ISSN 1592-7415 **eISSN 2282-8214**

Volume 32 – 2017

RATIO MATHEMATICA

Journal of Mathematics, Statistics, and Applications

Publisher

A.P.A.V.
Accademia Piceno – Aprutina dei Velati in Teramo
www.eiris.it – www.apav.it
apavsegreteria@gmail.com

RATIO MATHEMATICA
Journal of Mathematics, Statistics, and Applications
ISSN 1592-7415 - eISSN 2282-8214

Ratio Mathematica is an **International, double peer-reviewed, open access journal, published every six months** (June-December).

The main topics of interest for Ratio Mathematica are:

Foundations of Mathematics: Epistemology of Mathematics, Critique of the Foundations of Mathematics, Elementary Mathematics from a higher point of view, Elementary Theory of Numbers, Foundations of Mathematics of Uncertain; Criticism of the Foundations of Computer Science.

Applications of Mathematics: Applications to Engineering and Economics, Applications for Management and Business Administration, Decision making in conditions of uncertainty, Theory of Games, Fuzzy Logic and Applications, Probability, Statistics and Applications.

New theories and applications: Algebraic hyperstructures and Applications, Discrete Mathematics and Applications, Fuzzy structures.

New theories and practices for dissemination and communication of mathematics: Communication of History and Fundations, Communication of Discrete Mathematics, Communication of Probability and Statistics, Communication with Media.

Publisher: **APAV** - Accademia Piceno-Aprutina dei Velati in Teramo
Tax Code 92036140678, **Vat Id. Number** 02184450688
Registered Office Address: via del Concilio, 24 - 65121 Pescara
Operational Office: via Chiarini, 191 - 65126 Pescara

Ratio Mathematica **ISSN**: 1592-7415
Vol. 32, 2017, pp. 3–19 **eISSN**: 2282-8214

Some improved mixed regression estimators and their comparison when disturbance terms follow Multivariate t-distribution

Manoj Kumar[1], Vikas Bist[2] and Man Inder Kumar[3,*].

[1]Department of Statistics, Panjab University, Chandigarh, India
mantiwa@gmail.com

[2]Department of Mathematics, Panjab University, Chandigarh, India
bistvikas@gmail.com

[3]Department of Statistics, Panjab University, Chandigarh, India
maninderbajarh@gmail.com

Received on: 27-05-2017. **Accepted** on: 15-06-2017. **Published** on: 30-06-2017

doi:10.23755/rm.v32i0.330

Abstract

The Mean square error matrices, bias vector and risk functions of proposed improved mixed regression estimators are obtained by employing the small disturbance approximation technique under the condition, when disturbance terms follows multivariate t-distribution. Further, the risk function criterion is used to examine the efficiency of proposed improved mixed regression estimators.

Keywords: Stochastic restrictions; Mixed regression estimator; Stein- rule estimator; Multivariate t-distribution etc.
2010 AMS subject classifications: 62J05

1 Introduction

When incomplete prior information is expressible in the form of set of linear stochastic restrictions on the coefficients in a linear regression model, the method of mixed regression for the estimation of regression coefficients provides asymptotically a more efficient estimator than the least squares method that ignores the prior restrictions.
Stemming from the philosophy of stein-rule in this paper we proposed two families of improved estimators for the regression coefficients and study their properties when disturbances have multivariate t-distribution. For multivariate t - distribution see, [12], [10] and [3]. In section 2, we discuss the framework and estimators. The properties of these estimators are presented in section 3 and the results are compared in section 4. Simulation Study is carried out to support theoretical finding in Section 5.

2 Model Specification and the Estimators

Let us postulate the linear regression model

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

Where, Y is a $n \times 1$ vector of dependent variables; X is a $n \times p$ column rank matrix of n-observations on p explanatory non-stochastic variables; β is a $p \times 1$ non-null vector of regression coefficient and U is a $n \times 1$ vector of disturbance following multivariate student t-distribution with probability density function as:

$$f\left(\frac{U}{v}, \sigma^2\right) = \frac{\gamma^{v/2}\Gamma\left(\frac{v+n}{2}\right)}{\pi^{\frac{n}{2}}\Gamma\left(\frac{v}{2}\right)}\sigma^{-n}\left[v + \frac{U'U}{\sigma^2}\right]^{-\frac{n+v}{2}} \tag{2}$$

Where, $v > 0, \sigma > 0$ are respectively the degree of freedom and dispersion parameters; the vector U has its error components $U_i \in (-\infty, \infty), i = 1, 2, ..., n$. Here the error vector U has mean vector $E(U) = 0$ for $v > 1$, variance-covariance matrix $E(U'U) = \sigma^2\left(\frac{v}{v-2}\right) I$, for $v > 2$, measure of skewness $\gamma_1 = 0$ and measure of kurtosis $\gamma_2 = \sigma^4\left(\frac{6}{v-4}\right) I$ for $v > 4$.
Let the stochastic restrictions on β in (1) be

$$r = R\beta + V \tag{3}$$

Where, r is a $J \times 1$ vector of known elements, R is a $J \times p$ full row rank matrix of known elements and V is a $J \times 1$ vector of distribution such that

$$E(V) = 0\,;\; E(V'V) = \Omega \tag{4}$$

Where, Ω is a $J \times J$ positive definite symmetric matrix of known elements.

Further, we assume that the errors associated with the stochastic restriction are independent with the distribution in model (1).
The ordinary least square (OLS) estimator of β that ignores the prior restrictions (3) is

$$b_o = (X'X)^{-1}X'Y \tag{5}$$

If we consider the prior information (3), then the mixed regression (MR) estimator of β is given by

$$b_{MR} = [X'X + S^2 R'\Omega^{-1}R]^{-1}[X'Y + S^2 R'\Omega^{-1}r] \tag{6}$$

Where ,

$$S^2 = \frac{1}{n-p}((Y - Xb)'(Y - Xb)) \tag{7}$$

The Stein-rule estimator of β is

$$b_s = \left[1 - k\frac{(Y - Xb)'(Y - Xb)}{b_o'(X'X)b_0}\right] b_o \tag{8}$$

Where, k is a positive scalar characterizing the estimator.

The Stein-Mixed Regression (SMR) estimator of β is given as

$$b_{SMR} = \left[X'X + \frac{1}{n-p}[(Y - Xb_S)'(Y - Xb_S)]R'\Omega^{-1}R\right]^{-1} \left[X'Y + \frac{1}{n-p}[(Y - Xb_S)'(Y - Xb_S)]R'\Omega^{-1}r\right] \tag{9}$$

The Mixed Stein-Regression (MSR) estimator of β is

$$b_{MSR} = \left[1 - k\frac{(Y - Xb_{MR})'(Y - Xb_{MR})}{b_{MR}'(X'X)b_{MR}}\right] b_{MR} \tag{10}$$

3 Properties of the Estimators

$$P_X = X(X'X)^{-1}X' \tag{11}$$

$$M = [I - P_X] \tag{12}$$

$$M_j = [P_X - jC^{-1}X\beta\beta'X'] \qquad j = 1, 2, . \tag{13}$$

$$N_j = [(X'X)^{-1} - jC^{-1}\beta\beta'] \qquad j = 1, 2, . \tag{14}$$

$$C = \beta'X'X\beta \tag{15}$$

$$\mu = (X'X)^{-1}R'\Omega^{-1}R(X'X)^{-1} \tag{16}$$

The OLS estimator defined in (5) is found to be unbiased if $v > 1$, with variance - covariance matrix and risk function given by

$$E[(b_0 - \beta)(b_0 - \beta)'] = \sigma^2 \left(\frac{v}{v-2}\right)(X'X)^{-1}; \quad v > 2 \tag{17}$$

$$Risk(b_o) = \sigma^2 \left(\frac{v}{v-2}\right) tr(X'X)^{-1}L; \quad v > 2 \tag{18}$$

Where, L is a positive definite symmetric loss matrix.

The properties of the MR estimator are same as the SMR estimator, so we consider only the SMR estimator and present the results in the form of following theorems.

Theorem 3.1. *The asymptotic expression for the bias vector, mean squared error matrix and risk function of SMR estimator, up to order $o(\sigma^4)$ of approximations are given as*

$$B(b_{SMR}) = 0 \tag{19}$$

$$M(b_{SMR}) = \sigma^2 \left(\frac{v}{v-2}\right)(X'X)^{-1} - \sigma^4 V_1; \quad v > 4 \tag{20}$$

Where,

$$V_1 = \Bigg[\left(1 - \frac{2}{n-p} - \frac{6}{v-4}\theta\right)\mu + \frac{6}{(v-4)(n-p)} \Big(\mu X'(I_n * M)X(X'X)^{-1} + (X'X)^{-1}X'(I_n * M)X\mu\Big)\Bigg] \tag{21}$$

$$\theta = \frac{trM(I_n * M)}{(n-p)^2} \tag{22}$$

$$Risk(b_{SMR}) = \sigma^2 \left(\frac{v}{v-2}\right) tr(X'X)^{-1}L - \sigma^4 trV_1 L \tag{23}$$

Proof 3.1: To employ small disturbances asymptotic approximations. Let us write model (1) as

$$Y = X\beta + \sigma\omega \quad (U = \sigma\omega) \tag{24}$$

So that the i.i.d. elements of ω have multivariate-t distribution with mean zero for $v > 1$, variance $\left(\frac{v}{v-2}\right)$, for $v > 2$, measure of skewness $\gamma_1 = 0$ and measure of kurtosis $\gamma_2 = \left(\frac{6}{v-4}\right)$ for $v > 4$.

Now, using (24) in (5), we find

$$b_0 = \beta + \sigma(X'X)^{-1}X'\omega \tag{25}$$

So that

$$Y - Xb_0 = \sigma M\omega \tag{26}$$

Where

$$M = [I_n - X(X'X)^{-1}X'] \tag{27}$$

Using (25), we find up to order $o(\sigma)$ of approximations.

$$\frac{1}{b_o'(X'X)b_0} = C^{-1}[1 - 2\sigma C^{-1}\beta' D'\omega] \tag{28}$$

Now, using (25), (26), and (28) in (8), we get up to order $o(\sigma^2)$ of approximations.

$$b_s - \beta = \sigma(X'X)^{-1}X'\omega - \sigma^2 k\omega' M\omega C^{-1}\beta \tag{29}$$

and for the same order of approximation, we have

$$Y - Xb_s = \sigma M\omega - \sigma^2 k\omega' M\omega C^{-1}X\beta \tag{30}$$

Thus, using (30) and (3) in (9), we get

$$b_{SMR} - \beta = \sigma h_1 + \sigma^2 h_2 + \sigma^3 h_3 + \sigma^4 h_4 \tag{31}$$

Here,

$$h_1 = (X'X)^{-1}X'\omega \tag{32}$$

$$h_2 = \left(\frac{\omega' M\omega}{n-p}\right)(X'X)^{-1}R'\Omega^{-1}V \tag{33}$$

$$h_3 = \left(\frac{\omega' M\omega}{T-G}\right)\mu X'\omega \tag{34}$$

$$h_4 = \left(\frac{\omega' M\omega}{n-p}\right)^2 [k^2(n-p)C^{-1}(X'X)^{-1}R'\Omega^{-1}V - \mu R'\Omega^{-1}V] \tag{35}$$

It is easy to see that

$$E(h_1) = E(h_2) = E(h_3) = E(h_4) = 0 \tag{36}$$

Utilizing (36) in (31), we obtain the result (19) of the Theorem 1.

Now using (31), we get

$$\begin{aligned}(b_{SMR}-\beta)(b_{SMR}-\beta)' = \sigma^2 h_1 h_1' + \sigma^3(h_1 h_2' + h_2 h_1') \\ + \sigma^4(h_1 h_3' + h_2 h_2' + h_3 h_1')\end{aligned} \tag{37}$$

Here,

$$E(h_1 h_1') = (X'X)^{-1} \tag{38}$$

$$E(h_1 h_2') = E(h_2 h_1') = 0 \tag{39}$$

$$E(h_1 h_3') = \frac{1}{n-p}\left[\frac{6}{v-4}(X'X)^{-1}X'(I_n * M)X\mu + (n-p)\mu\right] \tag{40}$$

$$E(h_2 h_2') = \left[\left(\frac{6}{v-4}\right)\theta + \left(\frac{n-p+2}{n-p}\right)\right]\mu \tag{41}$$

Utilizing (38), (39), (40) and (41) in (37), we obtain the result (20) of the **Theorem 1**.

$$Risk(b_{SMR}) = trM(b_{SMR})L \tag{42}$$

Thus, result (23) of the Theorem 1 follows from (42).

Theorem 3.2. *The asymptotic expression for bias vector, mean squared error matrix and risk function of MSR estimator, up to order $o(\sigma^4)$ of approximations are given as*

$$\begin{aligned}B(b_{MSR}) = -\sigma^2\frac{kv(n-p)}{v-2}C^{-1}\beta + \sigma^4\Big[\frac{6k}{v-4}C^{-2} \\ \Big((trM_4(I_n * M))I + 2(X'X)^{-1}X'(I_n * M)X - C\theta\mu(X'X)\Big)\beta \\ + kC^{-2}\Big((n-p)(p-2)I - \frac{n-p+2}{n-p}C\mu(X'X)\Big)\beta\Big]\end{aligned} \tag{43}$$

Where $*$ denotes Hadamard product.

$$M(b_{MSR}) = \sigma^2 \left(\frac{v}{v-2}\right)(X'X)^{-1} - \sigma^4 \Big[V_1 + \frac{12k}{v-4}C^{-1} \Big[(X'X)^{-1}X'(I_n * M)X(X'X)^{-1} - C^{-1}\Big((X'X)^{-1}X'(I_n * M)X\beta\beta' + \beta\beta'X'(I_n * M)X(X'X)^{-1} + \left(\frac{k}{2}\right)(trM(I_n * M))\beta\beta'\Big)\Big] + 2k(n-p)N_{(2+\frac{k}{2}(n-p+2))}\Big] \quad (44)$$

$$Risk(b_{MSR}) = \sigma^2 \left(\frac{v}{v-2}\right) tr(X'X)^{-1}L - \sigma^4 \Big[trV_1L + 12\frac{k}{v-4}C^{-1}\Big(tr(X'X)^{-1}X'(I_n * M)X(X'X)^{-1}L - C^{-1}\Big(2\beta'X'(I_n * M)X(X'X)^{-1}L\beta + \frac{k}{2}(trM(I_n * M))\beta'L\beta\Big)\Big) + 2k(n-p)trN_{(2+\frac{k}{2}(n-p+2))}L\Big] \quad (45)$$

Proof 3.2: Using (3), (24) and (26) in (6), we obtain up to order $o(\sigma^2)$ of approximations.

$$b_{MR} = \beta + \sigma(X'X)^{-1}X'\omega + \sigma^2\left(\frac{\omega'M\omega}{n-p}\right)(X'X)^{-1}R'\Omega^{-1}V \quad (46)$$

Thus, for the same order of approximation, we have

$$\frac{1}{b'_{MR}(X'X)b_{MR}} = C^{-1}\Big[1 - 2\sigma C^{-1}\beta'X'\omega - \sigma^2 C^{-1}\Big(\frac{2}{n-p}\omega'M\omega V'\Omega^{-1}R\beta + \omega'M_D\omega\Big)\Big] \quad (47)$$

Using (46), we get up to order $o(\sigma^2)$of approximations.

$$Y - Xb_{MR} = \sigma M\omega - \sigma^2\left(\frac{\omega'M\omega}{n-p}\right)X(X'X)^{-1}R'\Omega^{-1}V \quad (48)$$

Using (46), (47) and (48) in (10), we obtain up to order $o(\sigma^4)$, we get

$$b_{MSR} - \beta = \sigma h_1^* + \sigma^2 h_2^* + \sigma^3 h_3^* + \sigma^4 h_4^* \quad (49)$$

Where

$$h_1^* = (X'X)^{-1}X'\omega \quad (50)$$

$$h_2^* = \left(\frac{\omega' M\omega}{n-p}\right)\left[(X'X)^{-1}R'\Omega^{-1}V - kC^{-1}\beta\right] \tag{51}$$

$$h_3^* = -\left(\frac{\omega' M\omega}{n-p}\right)\left[\mu + k(n-p)C^{-1}N_2)\right]X'\omega \tag{52}$$

$$\begin{aligned} h_4^* = k(\omega' M\omega)C^{-2}\Big(&\frac{2}{n-p}\omega' M\omega\beta\beta' R'\Omega^{-1}V + \omega' M_4\omega\beta \\ &+ 2(X'X)^{-1}X'\omega\omega' X\beta\Big) - \left(\frac{\omega' M\omega}{n-p}\right)^2 \\ &\left[\mu R'\Omega^{-1}V + kC^{-1}(\beta' V'\Omega^{-1}R + (n-p)I)(X'X)^{-1}R'\Omega^{-1}V\right] \end{aligned} \tag{53}$$

Here, it is easy to verify that

$$E(h_1^*) = 0 \tag{54}$$

$$E(h_2^*) = -k(n-p)C^{-1}\beta \tag{55}$$

$$E(h_3^*) = 0 \tag{56}$$

$$\begin{aligned} E(h_4^*) = \frac{6k}{v-4}C^{-2}\Big[&(trM_4(I_n * M))I + 2(X'X)^{-1}X'(I_n * M)X \\ - C\theta\mu(X'X)\Big]\beta + kC^{-2}\Big[&(n-p)(p-2)I - \left(\frac{n-p+2}{n-p}\right)C\mu(X'X)\Big]\beta \end{aligned} \tag{57}$$

Utilizing (54), (55), (56) and (57) in (53), we obtain the result (43) of the Theorem 2.

Now, using (53) we get

$$\begin{aligned} (b_{MSR} - \beta)(b_{MSR} - \beta)' = \sigma^2 h_1^* h_1^{*'} + \sigma^3(h_1^* h_2^{*'} + h_2^* h_1^{*'}) \\ + \sigma^4(h_1^* h_1^{*'} + h_2^* h_2^{*'} + h_3^* h_1^{*'}) \end{aligned} \tag{58}$$

Here, we see that

$$E(h_1^* h_1^{*'}) = \left(\frac{v}{v-2}\right)(X'X)^{-1} \tag{59}$$

$$E(h_1^* h_2^{*'}) = 0 \tag{60}$$

$$\begin{aligned} E(h_1^* h_3^{*'}) = \frac{6}{(v-4)(n-p)}\Big[&(X'X)^{-1}X'(I_n * M)X\mu \\ &+ k(n-p)C^{-1}(X'X)^{-1}X'(I_n * M)XN_2\Big] - \mu - k(n-p)C^{-1}N_2 \end{aligned} \tag{61}$$

$$E(h_2^* h_2^{*'}) = \mu \left[\frac{6}{v-4}\theta + \left(\frac{n-p+2}{n-p} \right) I \right] \\ + k^2 C^{-2} \beta\beta' \Big[\frac{6}{v-4} trM(I_n * M) + (n-p)(n-p+2) \Big] \quad (62)$$

Utilizing (59), (60), (61) and (62) in (58), we obtain the result (44) of the Theorem 2. Similarly, we can obtain the result (45) of the **Theorem 2**.

4 Comparison of the Estimators

4.1 The comparison the risk functions of OLS and SMR estimators

On comparison the risk functions of OLS and SMR estimators. We observe that up to order $o(\sigma^2)$ of approximations, both the estimators have same risk and for higher order of approximation, we see that

$$Risk(b_0) - Risk(b_{SMR}) = \\ \sigma^4 \Big[\frac{6}{v-2} \Big(\frac{2}{n-p} tr(X'X)^{-1} X'(I_n * M) X \mu L - \theta tr \mu L \Big) + \Big(\frac{n-p-2}{n-p} \Big) tr \mu L \Big] \quad (63)$$

If we choose $L = (X'X)$, then expression (63) becomes

$$Risk(b_0) - Risk(b_{SMR}) \\ = \sigma^4 \Big[\frac{6}{v-2} \Big(\frac{2}{n-p} tr(X'X)^{-1} X'(I_n * M) X (X'X)^{-1} R' \Omega^{-1} R \\ - \theta tr(X'X)^{-1} R' \Omega^{-1} R \Big) + \left(\frac{n-p-2}{n-p} \right) tr(X'X)^{-1} R' \Omega^{-1} R \Big] \quad (64)$$

Since, the expression (64) is positive semi-definite, so b_{SMR} dominates b_0 and as $v \to \infty$, expression (64) reduces to

$$Risk(b_0) - Risk(b_{SMR}) = \sigma^4 \left(\frac{n-p-2}{n-p} \right) tr(X'X)^{-1} R' \Omega^{-1} R \quad (65)$$

Which is positive semi-definite. Thus, b_{SMR} dominates b_0, so long as $n - p > 2$.

4.2 The comparison the risk functions of OLS and MSR estimators

On comparison the risks of OLS and MSR, we see that up to order $o(\sigma^2)$ of approximations, both the estimators have same risk and for higher order of approximations, we find that b_{MSR} dominates b_o so long as (65) is positive semi-definite and if we choose k to satisfy,

$$0 < k < \frac{2(n-p)}{T} \frac{C}{\beta' A \beta} \Big[tr(X'X)^{-1}L - 2C^{-1}\beta' L \beta + \frac{6}{(n-p)(v-4)} \\ \Big(tr(X'X)^{-1}X'(I_n * M)X(X'X)^{-1}L - 2C^{-1}\beta' X'(I_n * M)X(X'X)^{-1}L\beta \Big) \Big] \tag{66}$$

Where

$$T = \left[\frac{6}{v-4}(trM(I_n * M)) + (n-p)(n-p+2) \right] \tag{67}$$

If we choose $L = (X'X)$, then the above condition of dominance becomes

$$0 < k < \frac{2(n-p)}{T} \Big[p - 2 + \frac{6}{(n-p)(v-4)} \Big(tr(X'X)^{-1}X'(I_n * M)X \\ - 2C^{-1}\beta' X'(I_n * M)X\beta \Big) \Big] \tag{68}$$

And as $v \to \infty$, condition (68) reduces to

$$0 < k < \frac{2}{n-p+2}(p-2); p > 2 \tag{69}$$

Which is well known condition of dominance of stein-rule estimator over the least squares estimator.

4.3 The comparison the risk functions of SMR and MSR estimators

On comparing the risk function associated with the estimators SMR and MSR respectively, we observe that the estimator MSR dominates the estimator SMR so long as (30)holds and as $v \to \infty$ and again by choosing $L = (X'X)$, the condition of dominance becomes (69).

5 Simulation Results

The proposed estimator b_{SMR} is more efficient than OLSE under given linear model. Although, theoretically the results are drawn in equation (65), the proposed Stein-mixed Regression (SMR) estimator b_{SMR} is more efficient than ordinary least square estimator b_0 under condition $n - p > 2$. In this section, we perform simulations for exact equation (65) under conditions $n-p > 2, n > p > j$ with sigma equal to one.

Each result is based on 100,000 simulations runs using MATLAB. The result shown for $n = 10, 11, 12, 13, 14, 15$ in Table 1, 2, 3 & 4. The main finding of our numerical evaluation is following:-

1. The simulation results strongly support the theoretical findings.
2. The simulation result also explains the strength keep on increasing as we go for large value of n, p and j.
3. The results are independent of value of sigma.
4. Hence, b_{SMR} is more efficient than b_0 under condition $n - p > 2$.
5. The simulation results also reveals that b_{MSR} is also more efficient over b_0 (as it also depends on (65) under condition at (69)).

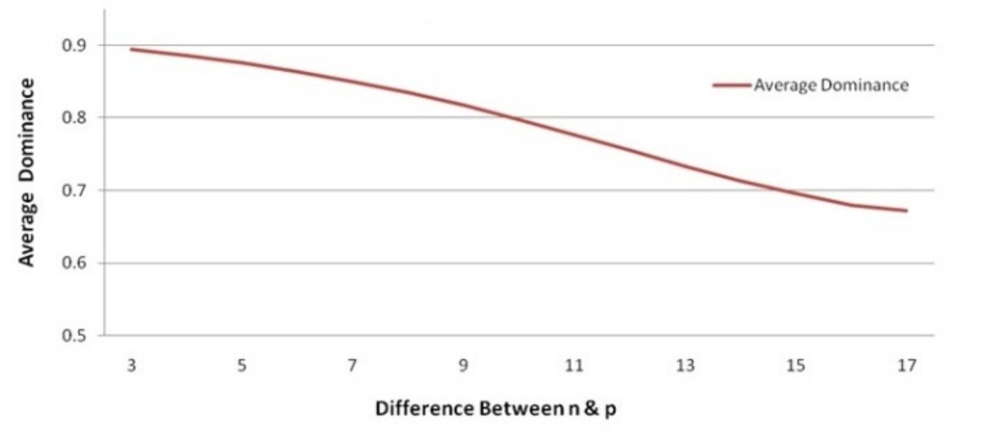

Figure 1: Average dominance condition for difference between n & p

Based on simulation study, the dominance of b_{SMR} has been proven over b_0 under certain set of conditions. Further, the behavior of dominance is studied for various combination of different values of n, p and j. The average dominance is derived based on probability for different combination of n, p and j; when $\sigma = 1$. The *figure 1* depicts average dominance keeps on decreasing with increase in gap

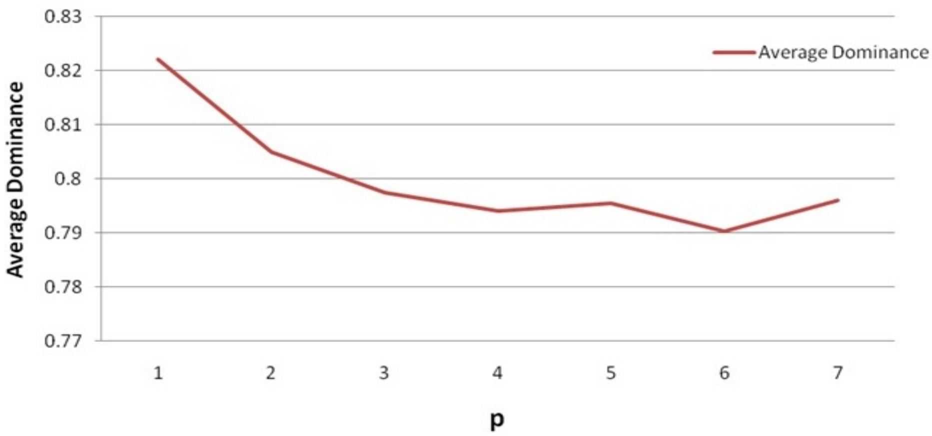

Figure 2: Dominance behavior for different values of p; when n=10 & j=5

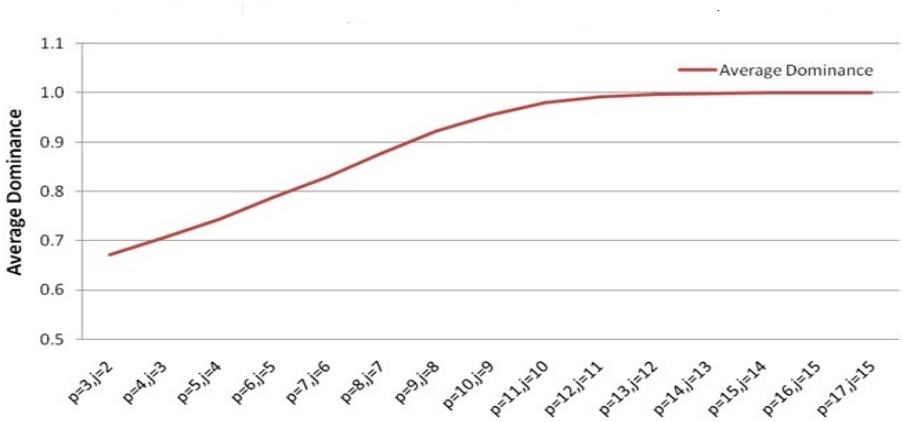

Figure 3: Average dominance condition for given value of p & j for n=20

between n and p. The *figure 2* also depicts a decreasing trend with increase in value of p, when $n = 10$ and $j = 5$. Similarly, *figure 3* shows the behavior of dominance condition for different value of p and j for fixed value of n equal to 20.

Table 1: Average Value of Dominance for different values of n and p for $j = 2$ $sigma = 1$.

j=2	n=10	n=11	n=12	n=13	n=14	n=15
p=3	0.67823	0.67492	0.67793	0.67688	0.67390	0.67539
p=4	0.67300	0.67079	0.66877	0.66903	0.66816	0.66654
p=5	0.67187	0.66562	0.66558	0.66668	0.66416	0.66691
p=6	0.66660	0.66642	0.66755	0.66509	0.66370	0.66186
p=7	0.66755	0.66497	0.66408	0.65994	0.66305	0.65982
p=8	-	0.66816	0.66611	0.66010	0.66261	0.66156
p=9	-	-	0.66861	0.66352	0.66143	0.66036

Remark: No value for dominance where $n - p \leqslant 2$.

Table 2: Average Value of Dominance for different values of n and p for $j = 3$ $sigma = 1$.

j=3	n=10	n=11	n=12	n=13	n=14	n=15
p= 4	0.70809	0.70806	0.70722	0.70725	0.70806	0.70611
p= 5	0.70570	0.70449	0.70581	0.70385	0.70381	0.70162
p= 6	0.70612	0.70319	0.70252	0.69998	0.70281	0.70205
p= 7	0.70651	0.70266	0.70281	0.70253	0.69810	0.69816
p= 8	-	0.70735	0.70167	0.70011	0.70183	0.69960
p= 9	-	-	0.70592	0.70486	0.70001	0.70028
p=10	-	-	-	0.70784	0.70426	0.70033
p=11	-	-	-	-	0.70692	0.70500

Remark: No value for dominance where $n - p \leqslant 2$.

Table 3: Average Value of Dominance for different values of n and p for $j = 4$ $sigma = 1$.

j=4	n=10	n=11	n=12	n=13	n=14	n=15
p= 5	0.74738	0.74891	0.74713	0.74538	0.74603	0.74668
p= 6	0.74694	0.74586	0.74771	0.74337	0.74335	0.74346
p= 7	0.75060	0.74841	0.74539	0.74412	0.74366	0.74381
p= 8	-	0.74740	0.74561	0.74405	0.74213	0.74607
p= 9	-	-	0.74752	0.74623	0.74547	0.74207
p=10	-	-	-	0.74601	0.74774	0.74091
p=11	-	-	-	-	0.74832	0.74426
p=12	-	-	-	-	-	0.74612

Remark: No value for dominance where $n - p \leqslant 2$.

Table 4: Average Value of Dominance for different values of n and p for $j = 5$ $sigma = 1$.

j=5	n=20	n=25	n=30	n=35	n=40	n=45	n=50
p = 5	0.78748	0.78793	0.78713	0.78373	0.78562	0.78336	0.78686
p=10	0.78359	0.78228	0.78284	0.78158	0.77898	0.77644	0.78005
p=15	0.78586	0.78143	0.77689	0.78072	0.77643	0.77454	0.77655
p=20	-	0.78350	0.78017	0.77976	0.77657	0.77576	0.77628
p=25	-	-	0.78453	0.78062	0.77884	0.77852	0.77563
p=30	-	-	-	0.78408	0.78044	0.77743	0.77703
p=35	-	-	-	-	0.78478	0.77772	0.77486
p=40	-	-	-	-	-	0.78208	0.77728

Remark: No value for dominance where $n - p \leqslant 2$.

References

[1] Chaturvedi, A. and Shukla, G., *Stein-rule estimation in linear models with non-scalar error covariance matrix*, Sankhya, Series B, 52, (1990), 293-304.

[2] Chaturvedi, A. ,Wan, A. T. K. and Singh, S. P., *Stein-rule restricted regres-*

sion estimator in a linear regression model with non-spherical disturbances, Communications in Statistics, Theory and Methods, 30, (2001), 55-68.

[3] Giles, A.J., *Pretesting for linear restriction in a regression model with spherically symmetric distributions*, Journal of Econometrics, 50, (1991), 377-398.

[4] Judge, G. G. and Bock, M. E., *The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics*, North Holland, Amsterdam, 1978.

[5] Kadane, J.B.,*Comparison of k-class Estimators When disturbance are small*, Econometrica, 39, (1971), 723-737.

[6] Ohtani, K. and Wan, A. T. K., *On the sampling performance of an improved Stein inequality restricted estimator*, Australian and New Zealand Journal of Statistics, 40, (1998), 181-187.

[7] Rao, C. R., *Linear Statistical Inference and Its Applications*, 2nd Edition. John Wiley,New York, 1973 .

[8] Shalabh and Wan, A. T. K., *Stein-rule estimation in mixed regression models*, Biometrical Journal, 42, (2000) 203-214.

[9] Sutradhar, B.C. and Ali, M.M., *Estimation of parameters of regression with a Multivariate t-error variable*, Communication Statistics - Theory and Methods, A 15, (1986), 429-450.

[10] Sutradhar, B.C., *Testing Linear Hypothesis with t - Error Variable*, Sankhya: The Indian Journal of Statistics, Series B (1960-2002), 50(2), (1988), 175-180.

[11] Theil, H., *Principles of Econometrics*, Vol. 1. Wiley, New York, 1971.

[12] Zellner, A., *Bayesian and non-Bayesian analysis of regression model with multivariate t-error terms*, Journal of the American Statistical Association, 71, (1976), 400-405.

Ratio Mathematica **ISSN**: 1592-7415
Vol. 32, 2017, pp. 21–35 **eISSN**: 2282-8214

On a Functional Equation Related to Information Theory

P. Nath[1], D.K. Singh[2*]

[1]Department of Mathematics, University of Delhi, Delhi - 110007, India
pnathmaths@gmail.com

[2]Department of Mathematics, Zakir Husain Delhi College (University of Delhi)
Jawaharlal Nehru Marg, Delhi - 110002, India
dhiraj426@rediffmail.com, dksingh@zh.du.ac.in

Received on: 28-05-2017. **Accepted** on: 15-06-2017. **Published** on: 30-06-2017

doi: 10.23755/rm.v32i0.332

Abstract

The main aim of this paper is to obtain the general solutions of the functional equation (1.3) without imposing any regularity condition on the mappings appearing in it. To do so, the general solutions of the functional equation (1.5), without imposing any regularity condition on the mappings appearing in it are needed. To meet this need, the general solutions of the functional equation (1.6) without imposing any regularity condition on a mapping appearing have to be investigated. One solution of (1.3) is useful in information theory. Thus, indeed, is the reason to consider (1.3).

Keywords: Functional equation; additive mapping; multiplicative mapping.
2010 AMS subject classifications: 39B22, 39B52, 94A15, 94A17.

*Corresponding author

1 Introduction

For $n = 1, 2, \ldots$, let $\Gamma_n = \{(p_1, \ldots, p_n) : 0 \leqslant p_i \leqslant 1, i = 1, \ldots, n; \sum_{i=1}^{n} p_i = 1\}$, denote the set of all discrete n-component complete probability distributions with non-negative elements. Let $I = \{x \in \mathbb{R} : 0 \leqslant x \leqslant 1\} = [0, 1]$, $\mathbb{R}$ denoting the set of all real numbers.

The axiomatic characterization of the non-additive entropy of degree α (see [2]) defined as

$$H_n^{\alpha}(p_1, \ldots, p_n) = (2^{1-\alpha} - 1)^{-1} \left(\sum_{i=1}^{n} p_i^{\alpha} - 1 \right), \quad \alpha \neq 1$$

leads to the study of the functional equation

$$\sum_{i=1}^{n} \sum_{j=1}^{m} f(p_i q_j) = \sum_{i=1}^{n} f(p_i) + \sum_{j=1}^{m} f(q_j) + \lambda \sum_{i=1}^{n} f(p_i) \sum_{j=1}^{m} f(q_j) \qquad (1.1)$$

in which $f : I \to \mathbb{R}$ is an unknown mapping, $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$, $\lambda \neq 0$, $\lambda \in \mathbb{R}$ and n, m being positive integers.

By a general solution of a functional equation, we mean a solution obtained without imposing any condition such as differentiability, continuity, continuity at a point, measurability, boundedness, monotonicity etc on a(the) mapping(s) appearing in the functional equation under consideration.

The general solutions of (1.1), for fixed integers $n \geqslant 3$, $m \geqslant 3$ and $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$ have been obtained in [5].

Losonczi [4] considered the functional equation

$$\sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij}(p_i q_j) = \sum_{i=1}^{n} h_i(p_i) + \sum_{j=1}^{m} k_j(q_j) + \lambda \sum_{i=1}^{n} h_i(p_i) \sum_{j=1}^{m} k_j(q_j) \qquad (1.2)$$

with $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$, $\lambda \neq 0$, $\lambda \in \mathbb{R}$, $f_{ij} : I \to \mathbb{R}$, $h_i : I \to \mathbb{R}$, $k_j : I \to \mathbb{R}$, $i = 1, \ldots, n$; $j = 1, \ldots, m$, as unknown mappings. He found the measurable (in the sense of Lebesgue) solutions of (1.2) for all $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$ by taking $n \geqslant 3$, $m \geqslant 3$ as fixed integers, in Theorem 6 on p-69 in [4]. For the last more than two decades, the general solutions of (1.2) for all $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$, $n \geqslant 3$, $m \geqslant 3$ being fixed integers, **are still not known so far**.

The main aim of this paper is to obtain the general solutions of the functional equation

$$\sum_{i=1}^{n} \sum_{j=1}^{m} h(p_i q_j) = \sum_{i=1}^{n} h(p_i) + \sum_{j=1}^{m} k_j(q_j) + \lambda \sum_{i=1}^{n} h(p_i) \sum_{j=1}^{m} k_j(q_j) \qquad (1.3)$$

which contains $m+1$ unknown real-valued mappings h and k_j ($j = 1, \ldots, m$), each defined on $I = [0, 1]$; $\lambda \in \mathbb{R}$, $\lambda \neq 0$ and $n \geqslant 3$, $m \geqslant 3$ being fixed integers. These general solutions have been obtained without making use of the difference operator D_i^r suggested on p-58 by Losonczi [4]. This paper is an improved version of the manuscript [9]. Nath and Singh [8] have also obtained the general solutions of

$$\sum_{i=1}^{n}\sum_{j=1}^{m} F(p_i q_j) = \sum_{i=1}^{n} G(p_i) + \sum_{j=1}^{m} H_j(q_j) + \lambda \sum_{i=1}^{n} G(p_i) \sum_{j=1}^{m} H_j(q_j)$$

with $F : I \to \mathbb{R}$, $G : I \to \mathbb{R}$, $H_j : I \to \mathbb{R}$, $j = 1, \ldots, m$; $\lambda \neq 0$, $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$, $n \geqslant 3$, $m \geqslant 3$ being fixed integers.

The functional equation (1.3) is a special case of (1.2). A particular case of (1.3) is the following

$$\sum_{i=1}^{n}\sum_{j=1}^{m} h(p_i q_j) = \sum_{i=1}^{n} h(p_i) + \sum_{j=1}^{m} k(q_j) + \lambda \sum_{i=1}^{n} h(p_i) \sum_{j=1}^{m} k(q_j)$$

in which $h : I \to \mathbb{R}$, $k : I \to \mathbb{R}$ and $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$. Nath and Singh [7] have obtained its general solution(s) for fixed integers $n \geqslant 3$, $m \geqslant 3$.

Let us define the mappings $f : I \to \mathbb{R}$ and $g_j : I \to \mathbb{R}$, $j = 1, \ldots, m$ as

$$f(x) = x + \lambda\, h(x); \qquad g_j(x) = x + \lambda\, k_j(x) \tag{1.4}$$

for all $x \in I$. Then (1.3) reduces to the Pexider type functional equation

$$\sum_{i=1}^{n}\sum_{j=1}^{m} f(p_i q_j) = \sum_{i=1}^{n} f(p_i) \sum_{j=1}^{m} g_j(q_j)\,. \tag{1.5}$$

We would like to mention that Kannappan and Sahoo [3] have obtained the general solutions of (1.3) and (1.5) on an open domain. In our case, the process of finding the general solutions of (1.5), for fixed integers $n \geqslant 3$, $m \geqslant 3$, needs determining the general solutions of the functional equation

$$\sum_{i=1}^{n}\sum_{j=1}^{m} \varphi(p_i q_j) = \sum_{i=1}^{n} \varphi(p_i) \sum_{j=1}^{m} \varphi(q_j) + m(n-1)\,\varphi(0) \sum_{i=1}^{n} \varphi(p_i) \tag{1.6}$$

where $\varphi : I \to \mathbb{R}$ and $n \geqslant 3$, $m \geqslant 3$ are fixed integers. This task has been accomplished in section 3. The corresponding general solutions of (1.5) and (1.3) have been investigated in sections 4 and 5 respectively. At the end of section 5, we have analysed the importance of the solutions of functional equation (1.3) from information-theoretic point of view. Section 2 contains some known definitions and results needed for the subsequent development of this paper.

2 Some preliminary results

In this section, we mention some known definitions and results.

A mapping $a : I \to \mathbb{R}$ is said to be additive on I or on the unit triangle $\Delta = \{(x, y) : 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1,\ 0 \leqslant x + y \leqslant 1\}$ if it satisfies the equation $a(x + y) = a(x) + a(y)$ for all $(x, y) \in \Delta$. A mapping $A : \mathbb{R} \to \mathbb{R}$ is said to be additive on $\mathbb{R}$ if it satisfies the equation $A(x + y) = A(x) + A(y)$ for all $x \in \mathbb{R}$, $y \in \mathbb{R}$. It is known [1] that if a mapping $a : I \to \mathbb{R}$ is additive on I, then it has a unique additive extension $A : \mathbb{R} \to \mathbb{R}$ in the sense that A is additive on $\mathbb{R}$ and $A(x) = a(x)$ for all $x \in I$.

A mapping $M : I \to \mathbb{R}$ is said to be multiplicative if $M(pq) = M(p)\,M(q)$ holds for all $p \in I$, $q \in I$.

Result 2.1. *[5] Let $n \geqslant 3$ be a fixed integer and c be a given constant. Suppose that a mapping $\psi : I \to \mathbb{R}$ satisfies the functional equation $\sum\limits_{i=1}^{n} \psi(p_i) = c$ for all $(p_1, \ldots, p_n) \in \Gamma_n$. Then there exists an additive mapping $b : \mathbb{R} \to \mathbb{R}$ such that $\psi(p) = b(p) - \dfrac{1}{n}\,b(1) + \dfrac{c}{n}$ for all $p \in I$.*

Result 2.2. *[4] Let d be a given real constant and $\psi_j : I \to \mathbb{R}$, $j = 1, \ldots, m$, be mappings which satisfy the functional equation $\sum\limits_{j=1}^{m} \psi_j(q_j) = d$ for all $(q_1, \ldots, q_m) \in \Gamma_m$, $m \geqslant 3$ being a fixed integer. Then there exists an additive mapping $a : \mathbb{R} \to \mathbb{R}$ and real constants c_j $(j = 1, \ldots, m)$ such that $\psi_j(p) = a(p) + c_j$ for all $p \in I$ with $a(1) + \sum\limits_{j=1}^{m} c_j = d$.*

3 The functional equation (1.6)

In this section, we prove:

Theorem 3.1. Let $n \geqslant 3$, $m \geqslant 3$ be fixed integers and $\varphi : I \to \mathbb{R}$ be a mapping which satisfies the functional equation *(1.6)* for all $(p_1, \ldots, p_n) \in \Gamma_n$ and $(q_1, \ldots, q_m) \in \Gamma_m$. Then φ is of the form

$$\varphi(p) = a(p) + \varphi(0) \tag{3.1}$$

where $a : \mathbb{R} \to \mathbb{R}$ is an additive mapping with

$$\begin{cases} \text{(i)} \quad a(1) = -\,nm\,\varphi(0) \quad \text{if} \;\; \varphi(1) + (n-1)\,\varphi(0) \neq 1 \\ \qquad\qquad \text{or} \\ \text{(ii)} \quad a(1) = 1 - n\,\varphi(0) \quad \text{if} \;\; \varphi(1) + (n-1)\,\varphi(0) = 1 \end{cases} \tag{3.2}$$

or

$$\varphi(p) = M(p) - B(p) \tag{3.3}$$

where $B : \mathbb{R} \to \mathbb{R}$ is an additive mapping with $B(1) = 0$ and $M : I \to \mathbb{R}$ is a multiplicative mapping which is not additive and $M(0) = 0$, $M(1) = 1$.

Proof. Let us put $p_1 = 1, p_2 = \ldots = p_n = 0$ in (1.6). We obtain

$$[\varphi(1) + (n-1)\,\varphi(0) - 1]\left[\sum_{j=1}^{m} \varphi(q_j) + m(n-1)\,\varphi(0)\right] = 0 \tag{3.4}$$

for all $(q_1, \ldots, q_m) \in \Gamma_m$. We divide our discussion into two cases.

Case 1. $\varphi(1) + (n-1)\,\varphi(0) \neq 1$.

In this case, (3.4) reduces to $\sum\limits_{j=1}^{m} \varphi(q_j) = -\,m(n-1)\,\varphi(0)$ for all $(q_1, \ldots, q_m) \in \Gamma_m$. By Result 2.1, there exists an additive mapping $a : \mathbb{R} \to \mathbb{R}$ such that φ is of the form (3.1) with $a(1)$ as in (3.2)(i). Thus, we have obtained the solution (3.1) satisfying (i) in (3.2).

Case 2. $\varphi(1) + (n-1)\,\varphi(0) - 1 = 0$.

Let us write (1.6) in the form

$$\sum_{j=1}^{m} \left\{ \sum_{i=1}^{n} \varphi(p_i q_j) - \varphi(q_j) \sum_{i=1}^{n} \varphi(p_i) - m(n-1)\varphi(0) q_j \sum_{i=1}^{n} \varphi(p_i) \right\} = 0\,.$$

By Result 2.1, there exists a mapping $A_1 : \Gamma_n \times \mathbb{R} \to \mathbb{R}$, additive in the second variable, such that

$$\begin{aligned} & \sum_{i=1}^{n} \varphi(p_i q) - \varphi(q) \sum_{i=1}^{n} \varphi(p_i) - m(n-1)\,\varphi(0)\, q \sum_{i=1}^{n} \varphi(p_i) \\ = \;& A_1(p_1, \ldots, p_n; q) - \varphi(0) \sum_{i=1}^{n} \varphi(p_i) + n\,\varphi(0) \end{aligned} \tag{3.5}$$

valid for all $(p_1, \ldots, p_n) \in \Gamma_n$ and $q \in I$ with

$$A_1(p_1, \ldots, p_n; 1) = m\,\varphi(0) \left[\sum_{i=1}^{n} \varphi(p_i) - n \right]. \tag{3.6}$$

Let $x \in I$ and $(r_1, \ldots, r_n) \in \Gamma_n$. Putting successively $q = x r_t$, $t = 1, \ldots, n$ in (3.5), adding the resulting n equations so obtained and then substituting the value of $\sum_{t=1}^{n} \varphi(x r_t)$ calculated from (3.5), we get the equation

$$\begin{aligned} & \sum_{i=1}^{n} \sum_{t=1}^{n} \varphi(x p_i r_t) - [\varphi(x) + m(n-1)\,\varphi(0)\, x - \varphi(0)] \\ & \quad \times \sum_{i=1}^{n} \varphi(p_i) \sum_{t=1}^{n} \varphi(r_t) - n^2\, \varphi(0) \\ & \qquad = A_1(p_1, \ldots, p_n; x) + m(n-1)\,\varphi(0)\, x \sum_{i=1}^{n} \varphi(p_i) \\ & \qquad\quad + A_1(r_1, \ldots, r_n; x) \sum_{i=1}^{n} \varphi(p_i)\,. \end{aligned} \tag{3.7}$$

The symmetry of the left hand side of (3.7), in p_i and r_t, $i = 1, \ldots, n$; $t = 1, \ldots, n$ gives rise to the equation

$$\begin{aligned} & [A_1(p_1, \ldots, p_n; x) + m(n-1)\,\varphi(0)\, x] \left[\sum_{t=1}^{n} \varphi(r_t) - 1 \right] \\ = \;& [A_1(r_1, \ldots, r_n; x) + m(n-1)\,\varphi(0)\, x] \left[\sum_{i=1}^{n} \varphi(p_i) - 1 \right]. \end{aligned} \tag{3.8}$$

Case 2.1. $\sum_{t=1}^{n} \varphi(r_t) - 1$ vanishes identically on Γ_n.

In this case, by Result 2.1, there exists an additive mapping $a : \mathbb{R} \to \mathbb{R}$ such that φ is of the form (3.1) but now $a(1)$ is as in (3.2)(ii).

Case 2.2. $\sum_{t=1}^{n} \varphi(r_t) - 1$ does not vanish identically on Γ_n.

Then, there exists a probability distribution $(r_1^*, \ldots, r_n^*) \in \Gamma_n$ such that

$$\left[\sum_{t=1}^{n} \varphi(r_t^*) - 1 \right] \neq 0 \,. \tag{3.9}$$

Setting $r_t = r_t^*$, $t = 1, \ldots, n$ in (3.8) and using (3.9), we obtain the equation

$$A_1(p_1, \ldots, p_n; x) = B(x) \left[\sum_{i=1}^{n} \varphi(p_i) - 1 \right] - m(n-1)\, \varphi(0)\, x \tag{3.10}$$

where $B : \mathbb{R} \to \mathbb{R}$ is defined as $B(x) = \left[\sum_{t=1}^{n} \varphi(r_t^*) - 1 \right]^{-1} [A_1(r_1^*, \ldots, r_n^*; x) + m(n-1)\, \varphi(0)\, x]$ for all $x \in \mathbb{R}$. It can be easily verified that $B : \mathbb{R} \to \mathbb{R}$ is an additive mapping with $B(1) = m\, \varphi(0)$. From (3.5), (3.10), $B(1) = m\, \varphi(0)$ and the additivity of $B : \mathbb{R} \to \mathbb{R}$, it follows that

$$\sum_{i=1}^{n} [M(p_i q) - M(q) M(p_i) + n(m-1)\, \varphi(0)\, M(q)\, p_i] = 0 \tag{3.11}$$

where $M : I \to \mathbb{R}$ is defined as

$$M(x) = \varphi(x) + B(x) + m(n-1)\varphi(0)x - \varphi(0) \tag{3.12}$$

for all $x \in I$. From (3.12), it is easy to see that $M(0) = 0$ as $B(0) = 0$. Applying Result 2.1 on (3.11), there exists a mapping $E : \mathbb{R} \times I \to \mathbb{R}$, additive in the first variable such that

$$M(pq) - M(p)M(q) + n(m-1)\, \varphi(0)\, M(q)\, p = E(p, q) - \frac{1}{n} E(1, q) \tag{3.13}$$

for all $p \in I$, $q \in I$. The substitution $p = 0$ in (3.13) and the use of $M(0) = 0$ gives $E(1, q) = 0$ for all $q \in I$. Consequently,

$$M(pq) - M(p)M(q) + n(m-1)\, \varphi(0)\, M(q)\, p = E(p, q) \tag{3.14}$$

for all $p \in I$, $q \in I$. Since $E(1,q) = 0$, therefore $E(1,1) = 0$. Now, putting $p = q = 1$ in equation (3.14), we obtain

$$M(1)[1 - M(1) + n(m-1)\varphi(0)] = 0\,. \tag{3.14a}$$

We prove that $M(1) \neq 0$. To the contrary, suppose that $M(1) = 0$. Putting $q = 1$ in (3.14) and using $M(1) = 0$, we get $M(p) = E(p,1)$ for all $p \in I$. So, M is additive on I. Also, if we put $x = 1$ in (3.12), use $M(1) = 0$ and $\varphi(1) + (n-1)\varphi(0) = 1$, we obtain $n(m-1)\varphi(0) = -1$. Now from (3.9), (3.12) and the additivity of M on I, we have $1 \neq \sum\limits_{t=1}^{n} \varphi(r_t^*) = 1$ a contradiction. Hence $M(1) \neq 0$. Now, from (3.14a), it follows that

$$M(1) - 1 = n(m-1)\varphi(0)\,. \tag{3.15}$$

Our next task is to prove that $M : I \to \mathbb{R}$, defined by (3.12), is not additive. To the contrary, suppose that M is additive. Now from (3.9), (3.12), the additivity of M on I and (3.15), we have

$$1 \neq \sum_{t=1}^{n} \varphi(r_t^*) = M(1) - n(m-1)\varphi(0) = 1$$

a contradiction. Hence $M : I \to \mathbb{R}$ is not additive.

Now we prove that, indeed, $M(1) - 1 = 0$. If possible, suppose $[M(1) - 1] \neq 0$. Putting $q = 1$ in (3.14) and using (3.15), we obtain

$$[M(1)p - M(p)] = [M(1) - 1]^{-1} E(p,1)$$

for all $p \in I$. Since $p \longmapsto E(p,1)$ is additive on I, it follows that $p \longmapsto M(1)p - M(p)$ must also be additive on I. But $p \longmapsto M(1)p$ is additive on I. Hence M is additive on I contradicting the fact that M is not additive. Hence $M(1) - 1 = 0$, that is, $M(1) = 1$.

Now, from (3.15), it follows that $\varphi(0) = 0$. Consequently, equation (3.14) reduces to the equation

$$M(pq) - M(p)\,M(q) = E(p,q) \tag{3.16}$$

for all $p \in I$, $q \in I$ and (3.12) reduces to (3.3) for all $p \in I$ with $B(1) = 0$.

The left hand side of (3.16) is symmetric in p and q. Hence $E(p,q) = E(q,p)$ for all $p \in I$, $q \in I$. Consequently, E is also additive on I in the second variable. We may assume that $E(p,\cdot)$ has been extended additively to the whole of $\mathbb{R}$.

Let $p \in I$, $q \in I$, $r \in I$. From (3.16), we have

$$\begin{aligned} E(pq,r) + M(r)\,E(p,q) &= M(pqr) - M(p)\,M(q)\,M(r) \\ &= E(qr,p) + M(p)\,E(q,r)\,. \end{aligned} \tag{3.17}$$

We prove that $E(p,q) = 0$ for all $p \in I$, $q \in I$. If possible, suppose there exists a $p^* \in I$ and a $q^* \in I$ such that $E(p^*,q^*) \neq 0$. Then, (3.17) gives

$$M(r) = [E(p^*,q^*)]^{-1}\{E(q^*r,p^*) + M(p^*)E(q^*,r) - E(p^*q^*,r)\}$$

from which it follows that M is additive on I contradicting the fact that M is not additive. Hence $E(p,q) = 0$ for all $p \in I$, $q \in I$. Now, from (3.16), it follows that $M(pq) = M(p)\,M(q)$ for all $p \in I$, $q \in I$. Thus, $M : I \to \mathbb{R}$ is a multiplicative mapping which is not additive and $M(0) = 0$, $M(1) = 1$. □

4 The functional equation (1.5)

In this section, we prove:

Theorem 4.1. Let $n \geqslant 3$, $m \geqslant 3$ be fixed integers and $f : I \to \mathbb{R}$, $g_j : I \to \mathbb{R}$, $j = 1, \ldots, m$ be mappings which satisfy the functional equation *(1.5)* for all

$(p_1, \ldots, p_n) \in \Gamma_n$ and $(q_1, \ldots, q_m) \in \Gamma_m$. Then, any general solution of *(1.5)*, for all $p \in I$, is of the form

$$\begin{cases} f(p) = b(p) \\ g_j \text{ any arbitrary real-valued mapping} \end{cases} \tag{4.1}$$

or

$$\begin{cases} f(p) = [f(1) + (n-1)\, f(0)]\, a(p) + f(0), \\ \qquad\qquad\qquad [f(1) + (n-1)\, f(0)] \neq 0 \\ g_j(p) = a(p) + A^*(p) + g_j(0) \end{cases} \tag{4.2}$$

or

$$\begin{cases} f(p) = f(1)[M(p) - B(p)]\,, \qquad\qquad f(1) \neq 0 \\ g_j(p) = M(p) - B(p) + A^*(p) + g_j(0) \end{cases} \tag{4.3}$$

where $b : \mathbb{R} \to \mathbb{R}$, $a : \mathbb{R} \to \mathbb{R}$, $A^* : \mathbb{R} \to \mathbb{R}$, $B : \mathbb{R} \to \mathbb{R}$ are additive mappings with

$$\begin{cases} \text{(i)} \quad b(1) = 0 \\ \text{(ii)} \quad B(1) = 0 \\ \text{(iii)} \quad a(1) = 1 - nf(0)[f(1) + (n-1)f(0)]^{-1} \\ \text{(iv)} \quad A^*(1) = -\sum_{j=1}^{m} g_j(0) + nm\, f(0)[f(1) + (n-1)f(0)]^{-1} \end{cases} \tag{4.4}$$

and $M : I \to \mathbb{R}$ is a multiplicative mapping which is not additive and $M(0) = 0$, $M(1) = 1$.

Proof. Put $p_1 = 1$, $p_2 = \ldots = p_n = 0$ in (1.5). We obtain

$$\sum_{j=1}^{m} [f(q_j) + (n-1)f(0)] = [f(1) + (n-1)\, f(0)] \sum_{j=1}^{m} g_j(q_j) \tag{4.5}$$

for all $(q_1, \ldots, q_m) \in \Gamma_m$.

Case 1. $f(1) + (n-1)\, f(0) = 0\,.$

Then, (4.5) reduces to the equation $\sum_{j=1}^{m} f(q_j) = -\, m(n-1)\, f(0)$. Put $q_1 = 1$, $q_2 = \ldots = q_m = 0$ in this equation and using the fact $f(1) + (n-1)\, f(0) = 0$, we have

$f(0) = 0 = f(1)$. Hence $\sum_{j=1}^{m} f(q_j) = 0$. By Result 2.1, there exists an additive mapping $b : \mathbb{R} \to \mathbb{R}$ such that $f(p) = b(p)$ with $b(1) = 0$. Consequently, for all $(p_1, \ldots, p_n) \in \Gamma_n$, $(q_1, \ldots, q_m) \in \Gamma_m$, it is easy to verify that $\sum_{i=1}^{n} \sum_{j=1}^{m} f(p_i q_j) = \sum_{i=1}^{n} f(p_i) = b(1) = 0$. Now, from (1.5), it follows that g_j ($j = 1, \ldots, m$) are, indeed, arbitrary real-valued mappings. Thus, we have obtained the solution (4.1) of (1.5) where $b(1)$ is given by (4.4)(i).

Case 2. $f(1) + (n-1)\, f(0) \neq 0$.

In this case, (4.5) can be written in the form

$$\sum_{j=1}^{m} \left\{ g_j(q_j) - [f(1) + (n-1)\, f(0)]^{-1} [f(q_j) + (n-1)\, f(0)] \right\} = 0\,. \tag{4.6}$$

By Result 2.2, there exists an additive mapping $A^* : \mathbb{R} \to \mathbb{R}$ such that

$$g_j(p) = [f(1) + (n-1)\, f(0)]^{-1} [f(p) - f(0)] + A^*(p) + g_j(0) \tag{4.7}$$

for $j = 1, \ldots, m$ with $A^*(1)$ given by (iv) in (4.4). The elimination of $\sum_{j=1}^{m} g_j(q_j)$ from equations (1.5) and (4.6) gives the equation

$$\begin{aligned} \sum_{i=1}^{n} \sum_{j=1}^{m} f(p_i q_j) &= [f(1) + (n-1)\, f(0)]^{-1} \sum_{i=1}^{n} f(p_i) \sum_{j=1}^{m} f(q_j) \\ &\quad + [f(1) + (n-1)\, f(0)]^{-1} m(n-1)\, f(0) \sum_{i=1}^{n} f(p_i) \end{aligned} \tag{4.8}$$

valid for all $(p_1, \ldots, p_n) \in \Gamma_n$ and $(q_1, \ldots, q_m) \in \Gamma_m$. Define a mapping $\varphi : I \to \mathbb{R}$ as

$$\varphi(x) = [f(1) + (n-1)\, f(0)]^{-1}\, f(x) \tag{4.9}$$

for all $x \in I$. Then (4.8) reduces to the functional equation (1.6) with $\varphi(1) + (n-1)\, \varphi(0) = 1$. So, we need to consider only those solutions of (1.6) which satisfy the requirement $\varphi(1) + (n-1)\, \varphi(0) = 1$.

The solutions (3.1) (with (3.2)(ii)) and (3.3) of (1.6) satisfy the condition $\varphi(1) + (n-1)\,\varphi(0) = 1$. Making use of (4.9), (4.7), (3.1) (with (3.2)(ii)) and (3.3), the solutions (4.2) and (4.3) can be obtained in which $B(1)$, $a(1)$ and $A^*(1)$ are given respectively by (ii), (iii) and (iv) in (4.4). $\square$

5 The functional equation (1.3)

In this section, we prove:

Theorem 5.1. Let $n \geqslant 3$, $m \geqslant 3$ be fixed integers and $h : I \to \mathbb{R}$, $k_j : I \to \mathbb{R}$, $j = 1, \ldots, m$ be mappings which satisfy the functional equation *(1.3)* for all $(p_1, \ldots, p_n) \in \Gamma_n$ and $(q_1, \ldots, q_m) \in \Gamma_m$ and $\lambda \neq 0$. Then, any general solution of *(1.3)*, for all $p \in I$, is of the form

$$\begin{cases} h(p) = \dfrac{1}{\lambda}\,[b(p) - p] \\ k_j \text{ any arbitrary real-valued mapping} \end{cases} \tag{5.1}$$

or

$$\begin{cases} h(p) = \dfrac{1}{\lambda}\big\{[\lambda(h(1) + (n-1)\,h(0)) + 1]\,a(p) - p\big\} + h(0), \\ \qquad\qquad\qquad [\lambda(h(1) + (n-1)\,h(0)) + 1] \neq 0 \\ k_j(p) = \dfrac{1}{\lambda}\big\{a(p) + A^*(p) - p\big\} + k_j(0) \end{cases} \tag{5.2}$$

or

$$\begin{cases} h(p) = \dfrac{1}{\lambda}\,\big\{[\lambda\,h(1) + 1][M(p) - B(p)] - p\big\}\,, \quad [\lambda\,h(1) + 1] \neq 0 \\ k_j(p) = \dfrac{1}{\lambda}\,\big\{M(p) - B(p) + A^*(p) - p\big\} + k_j(0) \end{cases} \tag{5.3}$$

where $b : \mathbb{R} \to \mathbb{R}$, $a : \mathbb{R} \to \mathbb{R}$, $A^* : \mathbb{R} \to \mathbb{R}$, $B : \mathbb{R} \to \mathbb{R}$ are additive mappings with

$$\begin{cases} \text{(i)} & b(1) = 0 \\ \text{(ii)} & B(1) = 0 \\ \text{(iii)} & a(1) = 1 - \lambda n\, h(0)[\lambda(h(1) + (n-1)h(0)) + 1]^{-1} \\ \text{(iv)} & A^*(1) = -\lambda \sum\limits_{j=1}^{m} k_j(0) + \lambda n m h(0)[\lambda(h(1) + (n-1)h(0)) + 1]^{-1} \end{cases} \tag{5.4}$$

and $M : I \to \mathbb{R}$ is a multiplicative mapping which is not additive and $M(0) = 0$, $M(1) = 1$.

Proof. From (1.4) and the solutions of the functional equation (1.5) i.e., (4.1), (4.2), (4.3) with (4.4); we obtain respectively the solutions (5.1), (5.2), (5.3) with (5.4); of the functional equation (1.3). The details are omitted. □

Remarks. The object of this remark is to point out the importance of various solutions of Theorem 5.1 from information-theoretic point of view.

1. The summand $\sum\limits_{i=1}^{n} h(p_i)$ of the mapping h appearing in (5.1) is independent of the probabilities $p_1, \ldots, p_n$. The solution (5.1) may be of some importance in information theory provided k_j is chosen as a suitable mapping of probability p, $p \in I$.

2. In solution (5.2), the summands $\sum\limits_{i=1}^{n} h(p_i)$ and $\sum\limits_{j=1}^{m} k_j(q_j)$ are independent of the probabilities $p_1, \ldots, p_n$ and $q_1, \ldots, q_m$ respectively. So, this solution does not seem to be of any relevance in information theory.

3. In solution (5.3)

$$\sum_{i=1}^{n} h(p_i) = \frac{1}{\lambda}\{\beta_1 \sum_{i=1}^{n} M(p_i) - 1\}$$

and

$$\sum_{j=1}^{m} k_j(q_j) = \frac{1}{\lambda}\Big\{\sum_{j=1}^{m} M(q_j) - 1\Big\} + \beta_2$$

where

$$\beta_1 = \lambda h(1) + 1$$

$$\beta_2 = nm\, h(0)[\lambda(h(1) + (n-1)h(0)) + 1]^{-1}\,.$$

If $\beta_1 = 1$ and $\beta_2 = 0$, then $\sum\limits_{i=1}^{n} h(p_i) = L_n^{\lambda}(p_1, \ldots, p_n)$ and $\sum\limits_{j=1}^{m} k_j(q_j) = L_m^{\lambda}(q_1, \ldots, q_m)$ where (see Nath and Singh [6])

$$L_t^{\lambda}(x_1, \ldots, x_t) = \frac{1}{\lambda}\Big[\sum_{i=1}^{t} M(x_i) - 1\Big]. \tag{5.5}$$

The non-additive measure of entropy $H_t^{\alpha}(x_1, \ldots, x_t) = (2^{1-\alpha} - 1)^{-1}(\sum\limits_{i=1}^{t} x_i^{\alpha} - 1)$, $\alpha \neq 1$, is a particular case of (5.5) when $\lambda = 2^{1-\alpha} - 1$, $\alpha > 0$, $\alpha \neq 1$ and $M : I \to \mathbb{R}$ is of the form $M(p) = p^{\alpha}$, $p \in I$, $\alpha \neq 1$, $\alpha > 0$, $0^{\alpha} := 0$, $1^{\alpha} := 1$.

References

[1] Z. Daróczy and L. Losonczi, *Über die Erweiterung der auf einer Punktmenge additiven Funktionen*, Publ. Math. (Debrecen) 14 (1967), 239–245.

[2] J. Havrda and F. Charvát, *Quantification method of classification process, concept of structural α-entropy*, Kybernetika (Prague) 3 (1967), 30–35.

[3] PL. Kannappan and P.K. Sahoo, *On the general solution of a functional equation connected to sum form information measures on open domain-VI*, Radovi Matematicki 8 (1992), 231–239.

[4] L. Losonczi, *Functional equations of sum form*, Publ. Math. (Debrecen) 32 (1985), 57–71.

[5] L. Losonczi and Gy. Maksa, *On some functional equations of the information theory*, Acta Math. Acad. Sci. Hung. 39 (1982), 73–82.

[6] P. Nath and D.K. Singh, *On a multiplicative type sum form functional equation and its role in information theory*, Applications of Mathematics 51(5) (2006), 495–516.

[7] P. Nath and D.K. Singh, *On a sum form functional equation and its role in information theory*, in Proc.: 8th National Conf. ISITA on "Information Technology: Setting Trends in Modern Era", (2008), 88–94.

[8] P. Nath and D.K. Singh, *On a functional equation containing an indexed family of unknown mappings*, Functional Equations in Mathematical Analysis (Edited by Themistocles Rassias and Janusz Brzdek), Springer 52 (2011), 671–687.

[9] P. Nath and D.K. Singh, *A sum form functional equation on a closed domain and its role in information theory*, arXiv:1508.05910v1 (2015), 1–18.

Ratio Mathematica **ISSN**: 1592-7415
Vol. 32, 2017, pp. 37–44 **eISSN**: 2282-8214

Quasi-Order Hypergroups determinated by $\mathscr{T}$-Hypergroups

Šárka Hošková-Mayerová
University of Defence,
Faculty of Military Technology,
Department of Mathematics and Physics,
Kounicova 65, 612 00 Brno, Czech Republic
sarka.mayerova@unob.cz

Received on: 15-03-2017. **Accepted** on: 02-05-2017. **Published** on: 30-06-2017

doi: 10.23755/rm.v32i0.333

Abstract

Quasi-order hypergroups were introduced by Jan Chvalina in 90s of the twentieth century. He proved that they form a subclass of the class of all hypergroups, i.e. structures with one associative hyperoperation fulfilling the reproduction axiom. In this paper a theorem which allows an easy description of all quasi-order hypergroups is presented. Moreover, some results concerning the relation of quasi-order and upper quasi-order hypergroups are given. Furthermore, the transformation hypergroups acting on tolerance spaces are defined and an example of them is mentioned.
Keywords. Quasi-order hypergroup, order hypergroup, tolerance relation, transformation semihypergroup, transformation hypergroup.
2010 AMS subject classifications: 20F60, 20N20.

The applications of mathematics in other disciplines, for example, in informatics, play a key role and they represent, in the last decades, one of the purpose, of the study of the experts of hyperstructures theory all over the world. Hyperstructure theory was introduced in 1934 by the French mathematician Marty [16], at the 8th Congress of Scandinavian Mathematicians, where he defined hypergroups based on the notion of hyperoperation, began to analyze their properties, and applied them to groups. In the following decades and nowadays, a number of different hyperstructures are widely studied from the theoretical point of view and for their applications to many subjects of pure and applied mathematics by many mathematicians. In a classical algebraic structure, the composition of two elements is an element, while in an algebraic hyperstructure, the composition of two elements is a set. Several books have been written on hyperstructure theory, see [6, 10, 17]. A recent book on hyperstructures [9] points out on their applications in rough set theory, cryptography, codes, automata, probability, geometry, lattices, binary relations, graphs and hypergraphs. Another book [10] is devoted especially to the study of hyperring theory. Several kinds of hyperrings are introduced and analyzed. The volume ends with an outline of applications in chemistry and physics, analyzing several special kinds of hyperstructures: hyperstructures and transposition hypergroups.

Hypergroups in the sense of Marty [16] form the largest class of multivalued systems that satisfies group-like axioms. It should be noted that various problems in non-commutative algebra lead to the introduction of algebraic systems in which the operations are not single-valued. The motivation for generalization of the notion of group resulted naturally from various problems in non-commutative algebra, another motivation for such an investigation came from geometry. Hypergroups have been used in algebra, geometry, convexity, automata theory, combinatorial problems of coloring, lattice theory, Boolean algebras, logic etc., over the years. Over the following decades, new and interesting results again appeared, but it is above all that a more luxuriant flourishing of hyperstructures has been seen in the last 20 years. It is not surprising that hypergroups as well as hypergroupoids, quasi-hypergroups, semihypergroups, hyperfields, hyper vector spaces, hyperlattices etc. have been studied.

The most complete bibliography up to 2002 can be found in the monograph of Pierguilio Corsini and Violeta Leoreanu: Applications of Hyperstructure Theory [9]. Another comprehensive list of literature is in monograph [17] and updated information is included in web site: http://aha.eled.duth.gr.

In the paper [2] special types of hypergroups, so called *quasi-order hypergroups* ($\mathbb{QOHG}$) and *order hypergroups* ($\mathbb{OHG}$), were introduced (cf. also [6, 9, 14, 5]).

First of all recall some basic terms and definitions. A *hyperoperation* "$\circ$" on a nonempty set H is a mapping from $H \times H$ to $\mathscr{P}^*(H)$ (all nonempty subsets of

H). The *hypergroupoid* is a pair $(H, \circ)$. The *quasi-hypergroup* is a hypergroupoid if the reproduction axiom ($a \circ H = H = H \circ a$ for any $a \in H$) is fulfilled. The quasi-hypergroup $(H, \circ)$ is called a *hypergroup* if moreover the hyperoperation "$\circ$" is associative $\big((a \circ b) \circ c = a \circ (b \circ c)$ for any $a, b, c \in H\big)$. Here for nonempty $A, B \subseteq H$ we put $A \circ B = \bigcup\limits_{a \in A,\, b \in B} a \circ b$. We denote $a \circ B$ instead of $\{a\} \circ B$, $a \in H$. See, e.g. [5, 6, 9, 7, 11].

Let $(H, *)$ and $(H', \star)$ be hypergroupoids. Then a mapping $f \colon H \longrightarrow H'$ is called *inclusion homomorphism* if it satisfies the condition:

$$f(x * y) \subseteq f(x) \star f(y) \qquad \text{for all pairs } x, y \in H.$$

Let X be a set and τ be a tolerance relation (i.e., reflexive and symmetric binary relation)—see [1]. Then the pair (X, τ) is a *tolerance space*.

Definition 1. The hypergroup $(H, \circ)$ is called a *quasi-order hypergroup*—cf. [2, 4, 9]—if

(i) $a \in a^3 = a^2$ for any $a \in H$, (1)

(ii) $a \circ b = a^2 \cup b^2$ for any $a, b \in H$. (2)

The hypergroup $(H, \circ)$ is called an *order hypergroup* if moreover

(iii) $a^2 = b^2$ implies $a = b$ for any $a, b \in H$. (3)

Using the methods occurring in [2, 4] the following theorem characterizing all quasi-order hypergroups can be proved. For the prove see [13]. (By a^2 we mean $a \circ a$.)

Theorem 1. *Let $(H, \circ)$ is a quasi-order hypergroup. Denote $K(a) = a^2$ for any $a \in H$. Then the system of sets $K(a)$ fulfills the following conditions:*

(i) $a \in K(a)$ *for any* $a \in H$, (4)

(ii) *if* $b \in K(a)$ *then* $K(b) \subseteq K(a)$. (5)

Conversely, if any system of subsets $K(a)$ of the set H, $a \in H$, fulfills (4) *and* (5)*, then there exists the only hyperoperation "$\circ$" on H such that $a \circ a = K(a)$ and $(H, \circ)$ is a quasi-order hypergroup.*

With respect to (3) the following corollary evidently holds:

Corollary 1. Under the assumptions of Theorem 1 the quasi-order hypergroup $(H, \circ)$ is an order hypergroup if and only if for $a \neq b$ there is $K(a) \neq K(b)$.

It is easy to show that if R is a quasi-ordering on a set H, then the pair $(H, \circ)$, where $a \circ b = R(a) \cup R(b)$, $a, b \in H$, is a quasi-order hypergroup. ($R(x)$ is an upper end of an element $x \in H$, i.e. the set $\{a \in H; a\, R\, x$ for each element $a \in H\}$). See e.g. [3, 11].

In [3] J. Chvalina introduced the concept of an upper quasi-order and upper order hypergroup.

Definition 2. A hypergroup $(H, \circ)$ is said an *upper quasi-order (upper order) hypergroup* if there exists a quasi-ordering (ordering) R such that $a \circ b = R(a) \cup R(b)$ for $a, b \in H$.

It can be shown that the classes of all quasi-order hypergroups and upper quasi-order hypergroups coincide. The same is true for the classes of all order hypergroups and upper order hypergroups. See [2, Theorem 1] or [9, Proposition 2 on p. 96]. These results can be easily proved using Theorem 1.

As we will need the above mentioned result of Prof. Jan Chvalina several times in this text we recall its formulation:

Proposition 1. [9, Proposition 2 on p. 96] A hypergroupoid $(H, \cdot)$ is a (quasi)-order hypergroup if and only if there exists a (quasi)-ordeg ρ on the set H, such that

$$\forall (a, b) \in H \times H, \qquad a \cdot b = \rho(a) \cup \rho(b),$$

where $\rho(a) = \{x \in H, a\, \rho\, x\}$.

Theorem 2. *Every quasi-order (order) hypergroup is an upper quasi-order (upper order) hypergroup.*

Proof. Let $(H, \circ)$ be a quasi-order hypergroup. Let us define a relation R on H as follows: $a\, R\, b$ iff $b \in a^2$ for each $a, b \in H$. Evidently $a\, R\, b$ iff $b^2 \subseteq a^2$. Then (4) and (5) imply that R is a quasi-ordering. Moreover, $R(a) = a^2$. Thus $a \circ b = a^2 \cup b^2 = R(a) \cup R(b)$.

If $(H, \circ)$ is even an order hypergroup, then by Corollary 1 there is $R(a) \neq R(b)$ for $a \neq b$, $a, b \in H$. Thus R is an ordering. $\square$

In [12] a more general concept of *subquasi-order hypergroup* is introduced. It is an open question whether a similar representation result as in Theorem 1 can be found for this generalization.

Now let us recall the definition of a transformation hypergroup. It was introduced in [15].

Recall first that *tolerance relation* is a reflective and symmetric relation on a set. This relation yields the concept of singularity in abstract mathematical expressions. This relation namely in connection with other structures moves corresponding mathematical theories to useful applications. Many publications are devoted to systematic investigation to tolerances on algebraic structures compatible with all operations of corresponding algebras. A certain survey of important results including valuable investment can be found in [1]. Tolerance space is a set endowed with a tolerance relation.

Definition 3. Let X be a set, $(G, \bullet)$ be a hypergroup and $\pi\colon X \times G \to X$ a mapping such that $\pi(\pi(x,t),s) \in \pi(x, t \bullet s)$, where

$$\pi(x, t \bullet s) = \{\pi(x,u); u \in t \bullet s)\}$$

for each $x \in X$, $s, t \in G$.

Then the triple $\mathscr{T} = (X, G, \pi)$ is called a *discrete transformation hypergroup* or an action of the hypergroup G on the phase set X. The mapping π is also usually said to be simply an action.

More generally, it is possible to consider the situation, where the phase space X is endowed with some additional structure. The interesting case is given in the following definition.

Definition 4. Let (X, τ) be a tolerance space (so called phase tolerance space), $(G, \bullet)$ be a semihypergroup (so called phase semihypergroup) and $\pi\colon X \times G \to X$ a mapping such that

(i) $\pi(\pi(x,t),s) \in \pi(x, t \bullet s)$, where $\pi(x, t \bullet s) = \{\pi(x,u); u \in t \bullet s)\}$ for each $x \in X$, $s, t \in G$;

(ii) if $x, y \in X$ are such that $x \,\tau\, y$, then $\pi(x,g) \,\tau\, \pi(y,g)$ holds for any $g \in G$.

Then $\mathscr{T} = (X, G, \pi)$ is a *transformation semihypergroup with phase tolerance space*. If, moreover, the pair $(G, \bullet)$ is a hypergroup (so called phase hypergroup), then the triple $\mathscr{T} = (X, G, \pi)$ is a *transformation hypergroup with phase tolerance space*.

In case the tolerance τ is trivial, i.e., $x \,\tau\, y$ if and only if $x = y$, the preceding definition coincides in fact with Definition 3.

Let us consider a discrete transformation hypergroup $\mathscr{T} = (X, G, \pi)$. It is possible to assign to each transformation hypergroup a commutative, extensive hypergroup with the support X (i.e., phase set of $\mathscr{T}$) as follows:

Let us define for arbitrary pair of elements $x, y \in X$ a binary hyperoperation $\odot\colon X \times X \to \mathscr{P}^*(X)$ in this way:

$$x \odot y = \pi(x, G) \cup \pi(y, G) \cup \{x, y\},$$

where $\pi(x, G) = \{\pi(x,u), u \in G\}$ and similarly for $\pi(y, G)$.

In the following we will need the next Lemma. The proof can be found in [4].

Lemma 1. *A hypergroupoid* $(H, \cdot)$ *such that* $a \in a^3 \subset a^2$, $a \cdot b = a^2 \cup b^2$ *for any* $a, b \in H$ *is a quasi-order hypergroup.*

Proposition 2. The pair $(X, \odot)$ is an extensive, commutative hypergroup.

The extensivity and commutativity of the hyperoperation is evident, so the pair $(X, \odot)$ is an extensive, commutative hypergroupoid. The conditions of Lemma 1 are satisfied too, so $(X, \odot)$ is an extensive, commutative hypergroup.

Remark 1. Even in case when $\mathscr{T}$ is a transformation semihypergroup we can assign a commutative, extensive hypergroup to this semihypergroup by the above described way.

The considered mapping is a functorial assignment which is described in the following way:

The above defined assignment determines a functor F from the category $\mathbb{DTH}$ of all discrete transformation hypergroups into the category $\mathbb{AH}$ of all commutative (abelian) hypergroups.

The functor $F = (F_O, F_m)$ (O-as objects, m-as morphisms) is defined as follows: $F_O(\mathscr{T}) = (X, \odot)$; $F_m(h_X, h_G) = h_X$. Consider $\mathscr{T}_i = (X_i, G_i, \pi_i) \in \mathbb{DTH}$ where $(X_i, \odot_i)$ are hypergroups, $i = 1, 2$ and the morhpisms $h_X \colon X_1 \to X_2$. Then

$$\begin{aligned} h_X(x \odot_1 y) = h\big(\pi(x, G) \cup \pi(y, G) \cup \{x, y\}\big) &= \big\{\pi(h_X(x), h_G(g), g \in G_1)\big\} \\ &\cup \big\{\pi(h_X(y), h_G(g), g \in G_1)\big\} \cup \big\{h_X(x), h_X(y)\big\} \subseteq \\ &\pi(h_X(x), G_2) \cup \pi(h_X(y), G_2) \cup \{h_X(x), h_X(y)\} \\ &= h_X(x) \odot_2 h_X(y) \end{aligned}$$

holds for all $x, y \in X_1$.

Theorem 3. *The pair* $(X, \odot)$ *is a quasi-order hypergroup determined by* $\mathscr{T}$*, shortly quasi-order* $\mathscr{T}$*-hypergroup.*

Proof. Let us define on $(X, \odot)$ a binary relation "ρ" as follows:

$$x \, \rho \, y \Leftrightarrow \exists \, u \in G \text{ such that, } \pi(x, u) = y \text{ or } x = y.$$

This relation is evidently reflexive. We will show, that it is transitive as well.
Let

1) $x = y$ and $y = z$, then $x = z$ and $x \, \rho \, z$,
2) $x = y$ and $\pi(y, v) = z$, then $\pi(x, v = z)$ so $x \, \rho \, z$,
3) $\pi(x, u) = y$ and $y = z$, then $\pi(x, u = z)$ so $x \, \rho \, z$,

4) $\pi(x,u) = y$ and $\pi(y,v) = z$, then $z = \pi(y,v) = \pi(\pi(x,u),v)$. From Definition 4 we have $\pi(\pi(x,u),v) \in \pi(x, u \odot v)$, thus there exists $w \in u \odot v$ such that $z = \pi(x,w)$. Hence we have $x \,\rho\, z$. So ρ is a quasi-order. It is well known that $\rho^2 = \rho \supset \operatorname{diag}(X)$, where $\operatorname{diag}(X) = \{(x,x); x \in X\}$.

Now for any pair of elements $x, y \in X$ we get $x \odot y = \rho(x) \cup \rho(y)$. So according Proposition 1 the pair $(X, \odot)$ is a quasi-order hypergroup determined by $\mathscr{T}$. Shortly it is a quasi-order $\mathscr{T}$-hypergroup. $\square$

References

[1] I. Chajda, Algebraic Theory of Tolerance Relations, The Palacky University Olomouc, Czech Republic (1991).

[2] J. Chvalina, Commutative hypergroups in the sence of Marthy, Proceeding of the Summer School (1994), 19–30, Horní Lipová, Czech Republic.

[3] J. Chvalina, Functional Graphs, Quasi-ordered Sets and Commutative Hypergroups, MU Brno, (1995), (in Czech).

[4] J. Chvalina, L. Chvalinová, State hypergroup of automata, Acta Mat. et. Inf. Univ. Ostraviensis 4 (1996), 105–120.

[5] J. Chvalina, Š. Hošková, Abelization of proximal H_ν-rings using graphs of good homomorphisms and diagonals of direct squares of hyperstructures, Internat. Congress on AHA 8 (Samothraki, Greece 2002), 147–158.

[6] P. Corsini, Prolegomena of Hypergroup Theory, Aviani Editore, 2nd edition, Tricesimo, (1993).

[7] P. Corsini, Hypergraphs and hypergroups, Algebra Universalis, 35, (1996), 548–555.

[8] P. Corsini, V. Leoreanu, Hypergroups and Binary Relations, Algebra Universalis, 43, (2000), 321–330.

[9] P. Corsini, V. Leoreanu, Applications of Hyperstructure Theory, Kluwer Academic Publishers, Dordrecht, Hardbound, (2003).

[10] B. Davvaz, V. Leoreanu-Fotea, Hyperring Theory and Applications, International Academic Publishers, Palm Harbor, USA, 2007.

[11] Š. Hošková, Abelization of weakly associative hyperstructures and their proximal modifications, Ph.D. thesis, Masaryk University Brno (2003), 73p.

[12] Š. Hošková, Upper order hypergroups as a reflective subcategory of subquasi-order hypergroups, Italian Journal of Pure and Applied Math., No. 20, 215–222, (2006).

[13] Š. Hošková, Representation of quasi-order hypergroups, Global Journal of Pure and Applied Mathematics (GJPAM), Vol.1 No. 2, 173–176, India, (2005).

[14] Š. Hošková, J. Chvalina, The unique square root condition for quasi-order hypergroups and the corresponding reflector for the category of all order-hypergroups, Proc. of 3^{th} International Conference Aplimat 2004, Bratislava, 471–476, Slovakia.

[15] Š. Hošková, J. Chvalina, Transformation tolerance hypergroup, Thai Mathematical Journal, 4(1), (2006), 63–72.

[16] F. Marty, : Sur une generalisation de la notion de groupe, Huitième Congr. math. Scan. (1934), Stockholm, 45–49.

[17] T. Vougiouklis, : Hyperstructures and Their Representations, Hadronic Press Monographs in Mathematics, Palm Harbor Florida 1994.

Ratio Mathematica **ISSN**: 1592-7415
Vol. 32, 2017, pp. 45–62 **eISSN**: 2282-8214

Soft Γ- Modules

Serkan Onar[1], Sanem Yavuz[2],Bayram Ali Ersoy[3]
[1]Yildiz Technical University, Department of Mathematics, Istanbul, Turkey
serkan10ar@gmail.com
[2]Yildiz Technical University, Department of Mathematics, Istanbul, Turkey
ssanemy@gmail.com
[3]Yildiz Technical University, Department of Mathematics, Istanbul, Turkey
ersoya@gmail.com

Received on: 14-04-2017. **Accepted** on: 28-06-2017. **Published** on: 30-06-2017

doi:10.23755/rm.v32i0.325

Abstract

In this paper, the definitions of soft Γ -module, soft Γ - module homomorphism and soft Γ -exactness are introduced with the aid of the concept of soft set theory introduced by Molodtsov. In the meantime, some of their properties and structural characteristics are investigated and discussed. Thereafter, several illustrative examples are given.
Keywords: soft set; soft module; Γ- ring; Γ -module; soft Γ- module; soft Γ -module homomorphism; soft Γ -module isomorphism; soft Γ -exactness.
2010 AMS subject classifications: 03E72; 08A72.

1 Introduction

In the real world, there are some various uncertainties but classical mathematical tools is not convenient for modeling these. Uncertain and unclear data which are contained by economy, engineering, environmental science, social science, medical science, business administration and many other fields are common. Although many diverse theories such as probability theory, soft set theory, intuitionistic fuzzy soft set theory and rough set theory are known and these present advantageous mathematical approaches for modeling of uncertainties, each of these theories have their inherent diffuculties.

In 1999, Molodstov [1] developed soft set theory which is considered a mathematical tool for working with uncertainties. Since the emergence of soft set theory attracts attention and especially recently works on the soft set theory is progressing rapidly. Maji et al. [2] described some operations on soft sets and these operations are used soft sets of decision making problems. Chen et al. [3] offered a new definition for decrease of parametrerization on soft sets. They made comparasion between this definition and concept of restriction of property in the rough set theory. In theory, Maji et al. [4] worked various operator on soft set. Kong et al. [5] developed definition of parametrerization reduction on soft set. Zou and Xiao suggested some approach of data analysis in case of insuffcent information on soft set. Jiang et al. presented a unique approach of the semantic decision making by means of ontological thinking and ontology-based soft sets.

Besides studies on classic module theory have continued and interesting results have been discovered recently. Macias Diaz et al. [6] studied on modules which are isomorphic to relatively divisible or pure submodules of each other. Abuhlail et al. [7] presented on topological lattices and their applications to module theory. On the other hand, Ameri et al. [8] investigated gamma module and Davvaz et al. [9] studied tensor product of gamma modules.

As for soft module theory, Sun et al. [10] presented the notion of soft set and soft module. Xiang [11] worked soft module theory. T.Shah et al. [12] defined the notion of primary decomposition in a soft ring and soft module, and derived some related properties. Erami et al. [13] gave the concept of a soft MV-module and soft MV- submodule. In these days, there are some studies reletad with soft sets. Ali et al. [14] investigated some new operations in soft set theory and Pei et al. [15] studied from soft sets to information systems. Xiao et al. [16] presented research on synthetically evaluating method for business competitive capacity based on soft set. Aktaş et al. [17] showed soft sets and soft groups and Acar et al. [18] also showed soft sets and soft rings.

The main purpose of this paper is to deal with algebraic structure of $\Gamma-$ module by applying soft set theory. The concept of soft $\Gamma-$ module is introduced, their characterization and algebraic properties are investigated by giving some several examples. In addition to this, soft $\Gamma-$ homomorphism , soft $\Gamma-$ isomorphism and their properties are introduced. After all, we make inferences that images of soft $\Gamma-$ homomorphisms and inverse images of soft $\Gamma-$ homomorphisms are soft $\Gamma-$ homomorphisms. Furthermore soft $\Gamma-$ exactness is investigated and illustrated with a related example.

2 Preliminaries

In this section, preliminary informations will be required to soft $\Gamma-$ modules. First of all we give basic concepts of soft set theory.

Definition 2.1. *[18] Let X denotes an initial universe set and E is a set of parameters. The power set of X is denoted by $P(X)$. A pair of (F,E) is called a soft set over X if and only if F is a mapping from E into the set of all subsets of X, i.e, F: $E \to P(X)$.*

Definition 2.2. *[18] Let (F,A) and (G,B) be two soft sets over a common universe X.*

i) *If $A \widetilde{\subseteq} B$ and $F(a) \widetilde{\subseteq} G(a)$ for all $a \in A$ then we say that (F,A) is a soft subset of (G,B), denoted by $(F,A) \widetilde{\subseteq} (G,B)$.*

ii) *If (F,A) is a soft subset of (G,B) and (G,B) is a soft subset of (F,A), then we say that (F,A) is a soft equal to (G,B), denoted by $(F,A) \widetilde{=} (G,B)$.*

Example 2.1. *Let $X = M_2(Z_3)$ denotes an initial universe set, i.e, 2×2 matrices with Z_3 terms and $E = \{\begin{bmatrix} \overline{0} & \overline{0} \\ \overline{0} & \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{1} & \overline{0} \\ \overline{0} & \overline{1} \end{bmatrix}\}$is a set of parameters. Then $F: E \to P(X)$ where $F(\begin{bmatrix} \overline{0} & \overline{0} \\ \overline{0} & \overline{0} \end{bmatrix}) = \{\begin{bmatrix} \overline{0} & \overline{1} \\ \overline{1} & \overline{1} \end{bmatrix}, \begin{bmatrix} \overline{2} & \overline{1} \\ \overline{0} & \overline{2} \end{bmatrix}\}, F(\begin{bmatrix} \overline{1} & \overline{0} \\ \overline{0} & \overline{1} \end{bmatrix}) = \{\begin{bmatrix} \overline{2} & \overline{0} \\ \overline{0} & \overline{2} \end{bmatrix}\}$. Clearly,$(F,E)$ is called a soft set over X.*

Definition 2.3. *[18] Let (F,A) and (G,B) be two soft sets over a common universe X. The intersection of (F,A) and (G,B) is defined as the soft set (H,C) satisfying the following conditions:*

i) *$C = A \cap B$.*

ii) *For all $c \in C$, $H(c) = F(c)$ or $G(c)$.*

In this case, we write $(F,A) \widetilde{\cap} (G,B) = (H,C)$.

Definition 2.4. *[18] Let (F,A) and (G,B) be two soft sets over a common universe X. The union of (F,A) and (G,B) is defined as the soft set (H,C) satisfying the following conditions:*

i) *$C = A \cup B$.*

ii) *For all* $c \in C$,

$$H(c) = \left\{ \begin{array}{ll} F(c) & \text{if } c \in A - B, \\ G(c) & \text{if } c \in B - A, \\ F(c) \cup G(c) & \text{if } c \in A \cap B. \end{array} \right\}$$

This is denoted by $(F, A) \widetilde{\cup} (G, B) = (H, C)$.

Definition 2.5. *[18] If* (F, A) *and* (G, B) *are two soft sets over a common universe* X*, then* (F, A) AND (G, B) *denoted by* $(F, A) \widetilde{\wedge} (G, B)$ *is defined as* $(F, A) \widetilde{\wedge} (G, B) = (H, C)$*, where* $C = A \times B$ *and* $H(x, y) = F(x) \cap G(y)$*, for all* $(x, y) \in C$.

Definition 2.6. *Let* $\{(F_i, A_i) : i \in I\}$ *be a non- empty family soft sets. The* $\wedge-$*intersection of a non-empty family soft sets is defined by* $(\psi, Y) = \widetilde{\wedge}_{i \in I}(F_i, A_i)$ *where* (ψ, Y) *is a soft set,* $Y = \prod\limits_{i \in I} A_i$ *and* $\psi(y) = \cap_{i \in I} F_i(y)$ *for every* $y = (y_i)_{i \in I} \in Y$.

Definition 2.7. *[18] If* (F, A) *and* (G, B) *are two soft sets over a common universe* X*, then* $(F, A)\, OR\, (G, B)$ *denoted by* $(F, A) \widetilde{\vee} (G, B)$ *is defined as* $(F, A) \widetilde{\vee} (G, B) = (H, C)$*, where* $C = A \times B$ *and* $H(x, y) = F(x) \cup G(y)$*, for all* $(x, y) \in C$.

Definition 2.8. *Let* $\{(F_i, A_i) : i \in I\}$ *be a non- empty family soft sets. The* $\vee-$*union of a non-empty family soft sets is defined by* $(\psi, Y) = \widetilde{\vee}_{i \in I}(F_i, A_i)$ *where* (ψ, Y) *is a soft set,* $Y = \prod\limits_{i \in I} A_i$ *and* $\psi(y) = \cup_{i \in I} F_i(y)$ *for every* $y = (y_i)_{i \in I} \in Y$.

On the other hand we will introduce modules and soft modules, then we will study some properties and theories of soft modules such as trivial soft module, whole soft module, the concepts of soft submodule and soft module homomorphisms.

Definition 2.9. *[10] Let* R *be a ring with identity.* M *is said to be a left* R*-module if left scalar multiplication* $\lambda : R \times M \to M$ *via* $(a, x) \mapsto ax$ *satisfying the axioms* $\forall r, r_1, r_2, 1 \in R; m, m_1, m_2 \in M$:

i) M *is an abelian group,*

ii) $r(m_1 + m_2) = rm_1 + rm_2, (r_1 + r_2)m = r_1 m + r_2 m,$

iii) $(r_1 r_2)m = r_1(r_2 m),$

iv) $1m = m.$

Left $R-$*module is denoted by* ${}_R M$ *or* M *for short. Similarly we can define right* R*- module and denote it by* M_R.

Example 2.2. *Let* $R = M_2(Z)$ *and* $M = \{\begin{bmatrix} a \\ b \end{bmatrix} | a, b \in Z\}$. *Then* M *is module on* R.

Definition 2.10. *[10] Let* M *be a left* R*- module,* A *be a any nonempty set and* (F, A) *is a soft set over* M. (F, A) *is said to be a soft module over* M *if and only if* $F(x)$ *is submodule over* M*, for all* $x \in A$.

Definition 2.11. *[10] Let* (F, A) *be a soft module over* M *then*

i) (F, A) *is said to be a trivial soft module over* M *if* $F(x) = 0$ *for all* $x \in A$*,where* 0 *is zero element of* M.

ii) (F, A) *is said to be an whole soft module over* M *if* $F(x) = M$ *for all* $x \in A$.

Proposition 2.1. *[10] Let* (F, A) *and* (G, B) *be two soft modules over* M.

1) $(F, A) \widetilde{\cap} (G, B)$ *is a soft module over* M.

2) $(F, A) \widetilde{\cup} (G, B)$ *is a soft module over* M *if* $A \cap B = \emptyset$.

Definition 2.12. *[10] If* (F, A) *and* (G, B) *be two soft modules over* M*, then* $(F, A) + (G, B)$ *is defined as* $(H, A \times B)$*, where* $H(x, y) = F(x) + G(y)$ *for all* $(x, y) \in A \times B$.

Proposition 2.2. *[10] Assume that* (F, A) *and* (G, B) *are two soft modules over* M*.Then* $(F, A) + (G, B)$ *is soft module over* M.

Definition 2.13. *[10] Suppose that* (F, A) *and* (G, B) *be two soft modules over* M *and* N *respectively. Then* $(F, A) \times (G, B) = (H, A \times B)$ *is defined as* $H(x, y) = F(x) \times G(y)$ *for all* $(x, y) \in A \times B$.

Proposition 2.3. *[10] Let* (F, A) *and* (G, B) *be two soft modules over* M *and* N *respectively. Then* $(F, A) \times (G, B)$ *is soft module over* $M \times N$.

Definition 2.14. *[10] Let* (F, A) *and* (G, B) *be two soft modules over* M*.Then* (G, B) *is soft submodule of* (F, A) *if*

i) $B \subset A$,

ii) $G(x) < F(x), \forall\, x \in B$.

This is denoted by $(G, B)\widetilde{<}(F, A)$.

Proposition 2.4. *[10] Let* (F, A) *and* (G, B) *be two soft modules over* M*.We say that* (G, B) *is soft submodule of* (F, A) *if* $G(x) \subseteq F(x), \forall x \in A$.

Definition 2.15. *[10] Assume that* $E = \{e\}$*, where* e *is unit of* A*.Then every soft module* (F, A) *over* M *at least have two soft modules* (F, A) *and* (F, E) *called trivial soft submodule.*

Proposition 2.5. *[10] Let (F, A) and (G, B) are two soft modules over M and (G, B) is soft submodule of (F, A). If $f : M \to N$ is a homomorphism of module, then $(f(F), A)$ and $(f(G), B)$ are all soft modules over N and $(f(G), B)$ is soft submodule of $(f(F), A)$.*

Definition 2.16. *[10] Let (F, A) and (G, B) be two soft modules over M and N respectively, $f : M \to N, g : A \to B$ be two functions. Then we say that (f, g) is a soft homomorphism if the following conditions are satisfied:*

i) *$f : M \to N$ is a homomorphism of module,*

ii) *$g : A \to B$ is a mapping,*

iii) *For all $x \in A$, $f(F(x)) = G(g(x))$.*

We say that (F, A) is a soft homomorphic to (G, B) which denoted by $(F, A) \tilde{-} (G, B)$. In this definition, if f is an isomorphism from M to N and g is a one-to-one mapping from A onto B, then we say that (F, A) is a soft isomorphism and that (F, A) is a soft isomorphic to (G, B), this is denoted by $(F, A) \tilde{\cong} (G, B)$.

Finally, we will define Γ- ring and Γ- module and their homomorphisms which are basic definitions for soft Γ- module.

Definition 2.17. *[8] Let R and Γ be additive abelian groups. Then we say that R is a Γ- ring if there exists a mapping:*

$. : R \times \Gamma \times R \to R$

$(r_1, \gamma, r_2) \to r_1 \gamma r_2$

such that for every $a, b, c \in R$ and $\alpha, \beta \in \Gamma$,the following hold:

i) $(a + b)\alpha c = a\alpha c + b\alpha c$,

ii) $a(\alpha + \beta)c = a\alpha c + a\beta c$,

iii) $a\alpha(b + c) = a\alpha b + a\alpha c$,

iv) $(a\alpha b)\beta c = a\alpha(b\beta c)$.

Definition 2.18. *[8] A subset A of a Γ- ring R is said to be a right ideal of R if A is an additive subgroup of R and $A\Gamma R \subseteq A$,where $A\Gamma R = \{a\alpha c |\ a \in A, \alpha \in \Gamma, r \in R\}$.*

A left ideal of R is defined in a similar way. If A is both right and left ideal, we say that A is an ideal of R.

Definition 2.19. *[8] If R and S are Γ- rings, then a pair (θ, φ) of maps from R into S is called a homomorphism from R into S if*

i) $\theta(x + y) = \theta(x) + \theta(y)$,

ii) *φ is an isomorphism on Γ,*

iii) $\theta(x\gamma y) = \theta(x)\varphi(\gamma)\theta(y)$.

Definition 2.20. *[8] Let R be a Γ- ring. A left Γ- module R is an additive abelian group M together with a mapping $. : R \times \Gamma \times M \to M$ such that for all $m, m_1, m_2 \in M$ and $\gamma, \gamma_1, \gamma_2 \in \Gamma, r, r_1, r_2 \in R$ the following hold:*

i) $r\gamma(m_1 + m_2) = r\gamma m_1 + r\gamma m_2,$
ii) $(r_1 + r_2)\gamma m = r_1\gamma m + r_2\gamma m,$
iii) $r(\gamma_1 + \gamma_2)m = r\gamma_1 m + r\gamma_2 m,$
iv) $r_1\gamma_1(r_2\gamma_2 m) = (r_1\gamma_1 r_2)\gamma_2 m.$
A right Γ *- module* R *is defined in analogous manner.*

Example 2.3. *Let* $R = \{\begin{bmatrix}\overline{k} & \overline{m}\end{bmatrix} |k, m \in Z_2\}$*, i.e,* 1×2 *matrices and* $\Gamma = \{\begin{bmatrix}\overline{0}\\ \overline{0}\end{bmatrix}, \begin{bmatrix}\overline{1}\\ \overline{0}\end{bmatrix}\} \in Z_2$*, where* Γ*is* 2×1 *matrices. Then we say that* R *is a* Γ*- ring. Similarly,* R *and* Γ *are same if we choose* $M = \{\begin{bmatrix}\overline{0} & \overline{0}\end{bmatrix}, \begin{bmatrix}\overline{1} & \overline{1}\end{bmatrix}\}$*, then* M *is* Γ *- module* R*.*

Definition 2.21. *[8] Presume that* $(M, +)$ *be an* Γ *- module* R *. A nonempty subset* N *of* $(M, +)$ *is said to be a left* Γ *- submodule* R *of* M *if* N *is a subgroup of* M *and* $R\Gamma N \subseteq N$*,where* $R\Gamma N = \{r\gamma n \,|\gamma \in \Gamma, r \in R, n \in N\}$*, that is for all* $n_1, n_2 \in N$ *and for all* $\gamma \in \Gamma, r \in R; n_1 - n_2 \in N$ *and* $r\gamma n \in N$*. In this case we write* $N \leq M$*.*

Example 2.4. *In previous example, let* $N = \{\begin{bmatrix}\overline{0} & \overline{0}\end{bmatrix}\} \subset M$ *and* $H : N \to P(M)$ *be a set valued function defined by* $H(a) = \{b \in M|R(a, \alpha, b) \Leftrightarrow a\alpha b \in \begin{bmatrix}\overline{0} & \overline{0}\end{bmatrix}\}$ *for all* $a \in N.H$ *is clear that* $H(\begin{bmatrix}\overline{0} & \overline{0}\end{bmatrix}) = (\begin{bmatrix}\overline{0} & \overline{0}\end{bmatrix})$ *is* Γ *- submodule* R *of* M*..*

Definition 2.22. *[8] Let* M *and* N *be arbitrary* Γ *- module* R *. A mapping* $f : M \to N$ *is a homomorphism of* Γ *- module* R *if for all* $x, y \in M$ *and* $\forall r \in R, \forall\gamma \in \Gamma$ *we have*
i) $f(x + y) = f(x) + f(y),$
ii) $f(r\gamma x) = r\gamma f(x).$
A homomorphism f is monomorphism if f is one-to-one and f is epimorphism if f is onto. f is called isomorphism if f is both monomorphism and epimorphism. We denote the set of all R_Γ*- homomorphisms from* M *into* N *by* $Hom_{R_\Gamma}(M, N)$ *or shortly by* $Hom_{R_\Gamma}(M, N)$*. In particular* $M = N$ *we denote* $Hom(M, M)$ *by* $End(M)$*.*

Definition 2.23. *[18] Let* M *be a nonempty set and a* $\Gamma-$*module. The pair* (F, A) *is a soft set over* M*. The set* $Supp(F, A) = \{x \in A : F(x) \neq \emptyset\}$ *is called a support of the soft set* (F, A)*. The soft set* (F, A) *is non-null if* $Supp(F, A) \neq \emptyset$*.*

3 Soft Γ- Modules

In this section, firstly we will define soft $\Gamma-$ modules, then we will give some operations on this modules.Throughout the section, M is a $\Gamma-$module.

Definition 3.1. *Let (F, A) be a non-null soft set over M. Then, (F, A) is said to be a soft $\Gamma-$module over M if $F(a)$ is a $\Gamma-$submodule M such that $F : A \to P(M)$, (i.e. $a \to F(a)$) for all $a \in A, y \in Supp(F, A)$.*

Example 3.1. *For consider the additively abelian groups $\mathbb{Z}_6 = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}\}$ and $\Gamma = \{\overline{0}, \overline{2}\}$. Let $. : \mathbb{Z}_6 \times \Gamma \times \mathbb{Z}_6 \to \mathbb{Z}_6, (m_1, \Gamma, m_2) = m_1 \Gamma m_2$. Hence $\mathbb{Z}_6$ is a $\Gamma-$ module. Let $A = \mathbb{Z}_6$ and $F : A \to P(M)$ be a set valued function defined by*

$$\begin{aligned} f(\overline{0}) &= f(\overline{2}) = f(\overline{4}) = \mathbb{Z}_6, \\ f(\overline{1}) &= f(\overline{3}) = f(\overline{5}) = \{\overline{0}, \overline{3}\} \end{aligned}$$

are $\Gamma-$submodule of $\mathbb{Z}_6$. Hence $(F, A,)$ is a soft $\Gamma-$module over $\mathbb{Z}_6$.

Example 3.2. *Let M is a $\Gamma-$ module and (F, A) be a soft set over $M.F : A \to P(M)$ is defined by $F(x) = \{y \in M |\ x\alpha y = 0\}$ for all $x \in A, \alpha \in \Gamma$. It is clear that (F, A) is a soft $\Gamma-$ module.*

Example 3.3. *For consider the additively abelian groups*

$$\begin{aligned} M &= R = \{\begin{bmatrix}\overline{0} & \overline{0}\end{bmatrix}, \begin{bmatrix}\overline{1} & \overline{0}\end{bmatrix}, \begin{bmatrix}\overline{0} & \overline{1}\end{bmatrix}, \begin{bmatrix}\overline{1} & \overline{1}\end{bmatrix}\} \subseteq (\mathbb{Z}_2)_{1\times 2} \\ \text{and } \Gamma &= \{\begin{bmatrix}\overline{0} \\ \overline{0}\end{bmatrix}, \begin{bmatrix}\overline{1} \\ \overline{0}\end{bmatrix}\} \subseteq (\mathbb{Z}_2)_{2\times 1} \end{aligned}$$

with addition defined as matrice addition. It is trivial that R is a $\Gamma-$ ring. Also M is a $\Gamma-$ module over R. Let $N = \{\begin{bmatrix}\overline{0} & \overline{0}\end{bmatrix}\} \subseteq M$ and $H : N \to P(M)$ be a set valued function defined by $H(a) = \{b \in M |\ R(a, \alpha, b) \leftrightarrow a\alpha b \in \begin{bmatrix}\overline{0} & \overline{0}\end{bmatrix}, \forall \alpha \in \Gamma\}$ for all $a \in N$. It is clear that $H(\begin{bmatrix}\overline{0} & \overline{0}\end{bmatrix}) = \{\begin{bmatrix}\overline{0} & \overline{0}\end{bmatrix}\}$ are sub $\Gamma-$ module of M. Hence (H, N) is soft $\Gamma-$ module of M.

Theorem 3.1. *Let (F, A) and (G, B) are two soft $\Gamma-$modules over M. Then $(F, A) \widetilde{\cap} (G, B)$ is a soft $\Gamma-$module over M if it is non-null.*

Proof. By definition, we have that $(F, A) \widetilde{\cap} (G, B) = (H, C)$ where $H(c) = F(x) \cap G(y)$ for all $c \in C$. We assume that (H, C) is a non-null soft set over M. If $c \in Supp(H, C)$, then $H(c) = F(x) \cap G(y) \neq \emptyset$. We know that $(F, A), (G, B)$ are both soft $\Gamma-$module over M, and so, the nonempty sets $F(x)$ and $G(y)$ are both $\Gamma-$submodule over M. Thus, $H(c)$ is a $\Gamma-$submodule over M for all $c \in Supp(H, C)$. In this position, $(H, C) = (F, A) \widetilde{\cap} (G, B)$ is a soft $\Gamma-$module over M. □

Theorem 3.2. *Let (F, A) and (G, B) are two soft $\Gamma-$modules over M. Then $(F, A) \widetilde{\cup} (G, B)$ is a soft $\Gamma-$module over M if $A \cap B = \emptyset$.*

Proof. By definition, we have that $(F,A)\,\widetilde{\cup}\,(G,B) = (H,C)$ where $H(c) = F(x) \cap G(y)$ for all $c \in C$. Note first that (H,C) is a non-null owing to the fact that $Supp(H,C) = Supp\,(F,A)\,\widetilde{\cup}\,(G,B)$. Suppose that $c \in Supp(H,C)$.Then $H(c) \neq \emptyset$ so we have $F(x), G(y) \neq \emptyset$. From the hypothesis $A \cap B = \emptyset$, we follow that $H(c) = F(x) \cap G(y)$. On the other hand $F(x) \cap G(y)$ is a soft $\Gamma-$module over M, we conclude that (H,C) is a soft $\Gamma-$module over M for all $c \in Supp(H,C)$. Consequently $(F,A)\,\widetilde{\cup}\,(G,B) = (H,C)$ is a soft $\Gamma-$module over M. $\square$

On the other hand, union of two soft $\Gamma-$ modules is not always soft $\Gamma-$ module. We will explain this situation with following example.

Example 3.4. *Let* $M = \mathbb{Z}_6 = \{\overline{0},\overline{1},\overline{2},\overline{3},\overline{4},\overline{5}\}$ *is a* $M_\Gamma-$*module,* $\Gamma = \{\overline{0},\overline{1}\}, A = \mathbb{Z}_2 = \{\overline{0},\overline{1}\}$ *and* $B = \mathbb{Z}_3 = \{\overline{0},\overline{1},\overline{2}\}$ *such that* $F(\overline{0}) = F(\overline{1}) = \{\overline{0},\overline{2},\overline{4}\}, G(\overline{0}) = G(\overline{1}) = G(\overline{2}) = \{\overline{0},\overline{3}\}$ $A \cap B = \{\overline{0},\overline{1}\}$. *If this condition is hold, then* $(F,A)\,\widetilde{\cup}\,(G,B)$ *is not a soft* $\Gamma-$*module over* M. *Indeed,* $H(\overline{1}) = \{\overline{0},\overline{2},\overline{3},\overline{4}\} \notin P(M)$.

Definition 3.2. *If* (F,A) *and* (G,B) *are two soft* $\Gamma-$*modules over* M*, then* (F,A) $AND\,(G,B)$ *denoted by* $(F,A)\,\widetilde{\wedge}\,(G,B)$ *is defined as* $(F,A)\,\widetilde{\wedge}\,(G,B) = (H,C)$*, where* $C = A \times B$ *and* $H\,(x,y) = F\,(x)\,\widetilde{\cap} G\,(y)$*, for all* $(x,y) \in C$.

Theorem 3.3. *Suppose that* (F,A) *and* (G,B) *are two soft* $\Gamma-$*modules over* M. *Then* $(F,A)\,\widetilde{\wedge}\,(G,B)$ *is soft* $\Gamma-$*module over* M *if it is non-null.*

Proof. Using definition, we have that $(F,A)\,\widetilde{\wedge}\,(G,B) = (H,C)$ where $C = A \times B$ and $H\,(x,y) = F\,(x)\,\widetilde{\cap} G\,(y)$, for all $(x,y) \in C$. Then the hypothesis, (H,C) is a non-null soft set over M. Since (H,C) is a non-null, $Supp\,(H,C) \neq \emptyset$ and so, for $(x,y) \in Supp\,(H,C)\,, H\,(x,y) = F\,(x)\,\widetilde{\cap} G\,(y) \neq \emptyset$. We assume that $t_1, t_2 \in F\,(x)\,\widetilde{\cap} G\,(y)$. In this position

i) If $t_1, t_2 \in F\,(x) = \{y : R(x,y)\}$ we have that $xt_1 \in A, xt_2 \in A$. This implies that $x(t_1 + t_2) \in A$.

ii) If $t_1, t_2 \in G\,(y) = \{y_1 : R(y,y_1)\}$ we have that $yt_1 \in B$, $yt_2 \in B$. This implies that $y(t_1 + t_2) \in B$.

Hence $F\,(x)\,\widetilde{\cap} G\,(y)$ is a $\Gamma-$ submodule. By the definition of soft $\Gamma-$ module, (F,A) and (G,B) are soft $\Gamma-$modules over M. $F\,(x)\,, G\,(y)$ are also $\Gamma-$ submodule over M. Furthermore $H\,(x,y) = F\,(x)\,\widetilde{\cap} G\,(y)$ is a $\Gamma-$ submodule over M for all $(x,y) \in (H,C) = (F,A)\,\widetilde{\wedge}\,(G,B)$. Hence $(F,A)\,\widetilde{\wedge}\,(G,B)$ is soft $\Gamma-$module over M. $\square$

Definition 3.3. *If* (F,A) *and* (G,B) *are two soft* $\Gamma-$*modules over* M*, then* (F,A) $OR\,(G,B)$ *denoted by* $(F,A)\,\widetilde{\vee}\,(G,B)$ *is defined as* $(F,A)\,\widetilde{\vee}\,(G,B) = (H,C)$*, where* $C = A \times B$ *and* $H\,(x,y) = F\,(x)\,\widetilde{\cup} G\,(y)$*, for all* $(x,y) \in C$.

Theorem 3.4. *Suppose that* (F,A) *and* (G,B) *are two soft* $\Gamma-$*modules over* M. *Then* $(F,A)\,\widetilde{\vee}\,(G,B)$ *is soft* $\Gamma-$*module over* M.

Proof. Using definition, we have that $(F,A)\,\widetilde{\vee}\,(G,B) = (H,C)$, where $C = A \times B$ and $H\left(x,y\right) = F\left(x\right)\widetilde{\cup}G\left(y\right)$, for all $(x,y) \in C$. Assume that $c \in Supp(H,C)$. Then $H(c) \neq \emptyset$ and so we have that $F(x) \neq \emptyset, G(y) \neq \emptyset$. By assumption, $F\left(x\right)\widetilde{\cup}G\left(y\right)$ is a soft $\Gamma-$ module of M for all $c \in Supp(H,C)$. Consequently $(F,A)\,\widetilde{\vee}\,(G,B) = (H,C)$ is a soft $\Gamma-$module over M. □

Definition 3.4. *Let (F,A) and (G,B) are two soft $\Gamma-$modules over M. Then $(F,A)\,\widetilde{+}\,(G,B) = (H, A \times B)$ is defined as $H(x,y) = F(x) + G(y)$ for all $(x,y) \in A \times B$.*

Theorem 3.5. *Suppose that (F,A) and (G,B) are two soft $\Gamma-$modules over M. Then $(F,A)\,\widetilde{+}\,(G,B)$ is soft $\Gamma-$module over M.*

Proof. By the definition we write $(F,A)\,\widetilde{+}\,(G,B) = (H, A \times B)$ and $H(x,y) = F(x) + G(y)$ for all $(x,y) \in A \times B$. Let $(x,y) \in Supp(H, A \times B)$.Then, $H(x,y) \neq \emptyset$ and so we have $F(x) \neq \emptyset$, $G(y) \neq \emptyset$. By taking into account, (F,A) and (G,B) are two soft $\Gamma-$modules over M, it follows that $F(x) + G(y)$ is a soft $\Gamma-$module over M for all $(x,y) \in Supp(H, A \times B)$. Hence $(F,A)\,\widetilde{+}\,(G,B)$ is soft $\Gamma-$module over M. □

Definition 3.5. *Let (F,A) and (G,B) are two soft $\Gamma-$modules over M. Then $(F,A)\,\widetilde{\times}\,(G,B) = (H, A \times B)$ is defined as $H(x,y) = F(x) \times G(y)$ for all $(x,y) \in A \times B$.*

Theorem 3.6. *Suppose that (F,A) and (G,B) are two soft $\Gamma-$modules over M. Then $(F,A)\,\widetilde{\times}\,(G,B)$ is soft $\Gamma-$module over M.*

Proof. By the definition we write $(F,A)\,\widetilde{\times}\,(G,B) = (H, A \times B)$ and $H(x,y) = F(x) \times G(y)$ for all $(x,y) \in A \times B$. Let $(x,y) \in Supp(H, A \times B)$. Then, $H(x,y) \neq \emptyset$ and so we have $F(x) \neq \emptyset$, $G(y) \neq \emptyset$. By taking into account, (F,A) and (G,B) are two soft $\Gamma-$modules over M, it follows that $F(x) \times G(y)$ is a soft $\Gamma-$module over M for all $(x,y) \in Supp(H, A\times B)$. Hence $(F,A)\,\widetilde{\times}\,(G,B)$ is soft $\Gamma-$module over M. □

Definition 3.6. *Let (F,A) and (G,B) are two soft $\Gamma-$modules over M. Then (G,B) is called a soft $\Gamma-$submodule of (F,A) if*

i) $B \subseteq A$,

ii) $\forall b \in Supp(G,B), g(b)$ *is a $\Gamma-$submodule of $F\left(b\right)$.*

This denoted by $(G,B) \subset (F,A)$. From the definition, it is easily deduced that if (G,B) is a soft $\Gamma-$submodule of (F,A), then $Supp(G,B) \subset Supp(F,A)$.

Theorem 3.7. *Let (F,A) and (G,B) be two soft $\Gamma-$modules over M and $(F,A)\widetilde{\subseteq}\left(G,B\right)$. Then $(G,B) \subset (F,A)$.*

Proof. Straight forward. □

Corolary 3.1. *Let* (F, A) *be a soft* $\Gamma-$*module over* M *and* $\{(F_i, A_i) : i \in I\}$ *be a nonempty family of soft* $\Gamma-$*submodules of* (F, A) *.Then,*

i) $\widetilde{\cap}_{i\in I}(F_i, A_i)$ *is a soft* $\Gamma-$*submodule of* (F, A) *if it is non-null.*

ii) $\widetilde{\cup}_{i\in I}(F_i, A_i)$ *is a soft* $\Gamma-$*submodule of* (F, A) *,* $if\ A_i \cap A_j = \emptyset$ *for all* $i, j \in I$ *and if it is non-null.*

iii) *If* $F_i(a_i) \subseteq F_j(a_j)$ *or* $F_j(a_j) \subsetneq F_i(a_i)$ *for all* $i, j \in I, a_i \in A_i$*, then* $\widetilde{\vee}_{i\in I}(F_i, A_i)$ *is a soft* $\Gamma-$*submodule of* $\widetilde{\vee}_{i\in I}(F, A)$ *.*

iv) $\widetilde{\wedge}_{i\in I}(F_i, A_i)$ *is a soft* $\Gamma-$*submodule of* $\widetilde{\wedge}_{i\in I}(F, A)$ *.*

v) *The cartesian product of the family* $\widetilde{\prod\limits_{i\in I}}(F_i, A_i)$ *is a soft* $\Gamma-$*submodule of* $\widetilde{\prod\limits_{i\in I}}(F, A)$ *.*

vi) $\widetilde{\sum}_{i\in I}(F_i, A_i)$ *is a soft* $\Gamma-$*submodule of* $\widetilde{\sum\limits_{i\in I}}(F, A)$ *.*

Proof. Similar to the proof of Theorems 3.5, 3.6, 3.9, 3.11, 3.13 and 3.15. □

4 Soft $\Gamma-$ Module Homomorphism

In this section, firstly we will define trivial and whole soft $\Gamma-$modules over $\Gamma-$module M, homomorphism of $\Gamma-$modules and their properties. Moreover we will study soft $\Gamma-$module homomorphism and soft $\Gamma-$module isomorphism. Throughout the section, M is a $\Gamma-$module.

Definition 4.1. *Let* (ρ, A) *and* (σ, B) *be two soft* $\Gamma-$*modules over* $\Gamma-$*module* M *and* $\Gamma-$*module* M_1 *respectively. Let* $f : M \to M_1$ *and* $g : A \to B$ *be two functions. The following conditions:*

i) f *is an epimorphism of* $\Gamma-$*module,*

ii) g *is a surjective mapping,*

iii) $f(\rho(y)) = \sigma(\rho(y))$ *for all* $y \in A$*,*

were satisfied by the pair (f, g)*, then* (f, g) *is called soft* $\Gamma-$*module homomorphism.*

If there exists a soft $\Gamma-$*module homomorphism between* (ρ, A) *and* (σ, B)*, we say that* (ρ, A) *is soft homomorphic to* (σ, B)*, and is denoted by* $(\rho, A) \sim (\sigma, B)$*. If there exists a soft* $\Gamma-$*module isomorphism between* (ρ, A) *and* (σ, B)*, we say that*(ρ, A) *is soft isomorphic to* (σ, B)*, and is denoted by* $(\rho, A)\widetilde{-}(\sigma, B)$*.*

Definition 4.2. *Let* (F, A) *be soft* $\Gamma-$*module over* M*.*

i) (F, A) *is called the trivial soft* $\Gamma-$*module over* M *if* $F(a) = \{0\}$ *for all* $a \in A$*.*

ii) *(F, A) is called the whole soft $\Gamma-$module over M if $F(a) = M$ for all $a \in A$.*

Definition 4.3. *Let M and M_1 be two $\Gamma-$modules and $m : M \to M_1$ a mapping of $\Gamma-$module. If (F, A) and (G, B) are soft sets over M and M_1 respectively, then*

i) *$(m(F), A)$ is a soft set over M_1 where $m(F) : A \to P(M_1)$, $m(F)(a) = m(F(a))$ for all $a \in A$.*

ii) *$(m^{-1}(G), B)$ is a soft set over M where $m^{-1}(G) : B \to P(M)$, $m^{-1}(G)(b) = m^{-1}(G(b))$ for all $b \in B$.*

Corolary 4.1. *Let $m : M \to M_1$ be an onto homomorphism of $\Gamma-$module. Then following statements can be given.*

i) *(F, A) be soft $\Gamma-$module over M, then $(m(F), A)$ is a soft $\Gamma-$module over $\Gamma-$module M_1.*

ii) *(G, B) be soft $\Gamma-$module over $\Gamma-$module M_1, then $(m^{-1}(G), B)$ is a soft $\Gamma-$module over M.*

Proof. i) Since (F, A) is a soft $\Gamma-$module over M, it is clear that $(m(F), A)$ is a non-null soft set over M_1. For every $y \in Supp(m(F), A)$ we have $m(F)(y) = m(F(y)) \neq \emptyset$. Hence $m(F(y))$ which is the onto homomorphic image of $\Gamma-$module $F(y)$ is a $\Gamma-$module of M_1 for all $y \in Supp(F(m), A)$. That is $(m(F), A)$ is a soft $\Gamma-$module over $\Gamma-$module M_1.

ii) It is easy to see that $Supp(m^{-1}(G), B) \subseteq Supp(G, B)$. By this way let $y \in Supp(m^{-1}(G), B)$.Then, $G(y) \neq \emptyset$. Hence $m^{-1}(G(y))$ which is homomorphic inverse image of $\Gamma-$module $G(y)$, is a soft $\Gamma-$module over M for all $y \in B$. $\square$

Theorem 4.1. *Let $m : M \to M_1$ be a homomorphism of $\Gamma-$module and (F, A), (G, B) be two soft $\Gamma-$modules over $\Gamma-$module M and $\Gamma-$module M_1 respectively. Then following statements can be given.*

i) *If $F(a) =$ ker (m) for all $a \in A$,then $(m(F), A)$ is the trivial soft $\Gamma-$module over M_1.*

ii) *If m is onto and (F, A) is whole, then $(m(F), A)$ is the whole soft $\Gamma-$module over M_1.*

iii) *If $G(b) = m(M)$ for all $b \in B$,then $(m^{-1}(G), B)$ is the whole soft $\Gamma-$module over M.*

iv) *If m is injective and (G, B) is trivial, then $(m^{-1}(G), B)$ is the trivial soft $\Gamma-$module over M.*

Proof. i) By using $F(a) =$ ker (m) for all $a \in A$. Then $m(F)(a) = m(F(a))$ $= \{0_{M_1}\}$ for all $a \in A$. Hence $(m(F), A)$ is soft $\Gamma-$module over M_1.

ii) Suppose that m is onto and (F, A) is whole. Then $F(a) = M$ for all $a \in A$ and so $m(F)(a) = m(F(a)) = m(M) = M_1$ for all $a \in A$. Hence $(m(F), A)$ is whole soft $\Gamma-$module over M_1.

iii) If we use hypothesis $G(b) = m(M)$ for all $b \in B$, we can write $m^{-1}(G)(b) = m^{-1}(G(b)) = m^{-1}(m(M)) = M$ for all $b \in B$. It is clear that, $(m^{-1}(G), B)$ is the whole soft $\Gamma-$module over M.

iv) Suppose that m is injective and (G, B) is trivial. Then, $G(b) = \{0\}$ for all $b \in B$,so $m^{-1}(G)(b) = m^{-1}(G(b)) = m^{-1}(\{0\}) = \ker m = \{0_M\}$ for all $b \in B$. Consequently, $(m^{-1}(G), B)$ is the trivial soft $\Gamma-$module over M. $\square$

Theorem 4.2. *Let* $m : M \to M_1$ *be a homomorphism of* $\Gamma-$*module and* (F, A)*,* (G, B) *be two soft* $\Gamma-$*modules over* M*. If* (G, B) *is soft* $\Gamma-$*submodule of* (F, A)*, then* $(m(G), B)$ *is soft* $\Gamma-$*submodule of* $(m(F), A)$*.*

Proof. Suppose that $y \in Supp\,(G, B)$. Then $y \in Supp\,(F, A)$ .We know that $B \subseteq A$ and $G(y)$ is a $\Gamma-$submodule $F(y)$ for all $y \in Supp\,(G, B)$. From the expression hypothesis m is a homomorphism, $m(G)(y) = m(G(y))$ is a $\Gamma-$submodule of $m(F)(y) = m(F(y))$ and therefore $(m(G), B)$ is soft $\Gamma-$submodule of $(m(F), A)$. $\square$

Theorem 4.3. *Let* $m : M \to M_1$ *be a homomorphism of* $\Gamma-$*module and* (F, A)*,* (G, B) *be two soft* $\Gamma-$*modules over* M*. If* (G, B) *is soft* $\Gamma-$*submodule of* (F, A)*, then* $(m^{-1}(G), B)$ *is soft* $\Gamma-$*submodule of* $(m^{-1}(F), A)$*.*

Proof. Let $y \in Supp(m^{-1}(G), B)$. $B \subseteq A$ and $G(y)$ is a $\Gamma-$submodule of $F(y)$ for all $y \in B$. Since m is a homomorphism, $m^{-1}(G)(y) = m^{-1}(G(y))$ is a $\Gamma-$submodule of $m^{-1}(G(y)) = m(G)(y)$ for all $y \in Supp(m^{-1}(G), B)$. Hence $(m^{-1}(G), B)$ is soft $\Gamma-$submodule of $(m^{-1}(F), A)$. $\square$

5 Soft $\Gamma-$ Exactness

In this section, we will introduce maximal and minimal soft $\Gamma-$submodules. Then, we will investigate short exact and exact sequence of $\Gamma-$modules. Finally, we will explain soft $\Gamma-$exactness and some their basic theories. Throughout this section M is $\Gamma-$module.

Definition 5.1. *Let* (F, A) *and* (G, B) *be two soft* $\Gamma-$*modules over* M *and* (G, B) *be soft* $\Gamma-$*submodule of* (F, A) *. We say* (G, B) *is maximal soft* $\Gamma-$*submodule of* (F, A) *if* $G(x)$ *is a maximal* $\Gamma-$*submodule of* $F(x)$ *for all* $x \in B$*. We say* (G, B) *is minimal soft* $\Gamma-$*submodule of* (F, A) *if* $G(x)$ *is a minimal* $\Gamma-$*submodule of* $F(x)$ *for all* $x \in B$*.*

Proposition 5.1. *Let* (F,A) *be a soft* $\Gamma-$*module over* M.

i) *If* $\{(G_i,B_i)\,|i\in I\}$ *is a nonempty family of maximal soft* $\Gamma-$*submodules of* (F,A), *then* $\bigcap_{i\in I}(G_i,B_i)$ *is maximal soft* $\Gamma-$*submodule of* (F,A).

ii) *If* $\{(G_i,B_i)\,|i\in I\}$ *is a nonempty family of minimal soft* $\Gamma-$*submodules of* (F,A), *then* $\sum_{i\in I}(G_i,B_i)$ *is minimal soft* $\Gamma-$*submodule of* (F,A).

Proof. straight forward. □

Corolary 5.1. *Let* (F,A) *be a soft* $\Gamma-$*module over* M *and* $f:M\to N$ *be a homomorphism if* $F(x)=\ker f$ *for all* $x\in A$, *then* $(f(F),A)$ *is the rivial soft* $\Gamma-$*module over* N. *Similarly, let* (F,A) *be an whole soft* $\Gamma-$*module over* M *and* $f:M\to N$ *be an epimorphism, then* $(f(F),A)$ *is a whole soft* $\Gamma-$*module over* N.

Definition 5.2. *The homomorphism sequence of* $\Gamma-$*modules* ... $\to M_{n-1}\to^{f_{n-1}} M_n\to^{f_n} M_{n+1}\to$ ... *is called exact sequence of* $\Gamma-$*modules if Im*$f_{n-1}=$ *Ker*f_n *for all* $n\in\mathbb{N}$ *and we call the exact sequence of* $\Gamma-$*modules form as* $0\to M_1\to^f M\to^g M_2\to 0$ *the short exact sequence of* $\Gamma-$*modules.*

Proposition 5.2. *Let* (F,A) *be a trivial soft* $\Gamma-$*module over* $\Gamma-$*module* M_1 *and* (G,B) *be a whole soft* $\Gamma-$*module over* $\Gamma-$*module* M_2 *if* $0\to M_1\to^f\to M\to^g\to M_2\to 0$ *is a short exact sequence, then* $0\to F(x)\to^{\widetilde{f}} M\to^{\widetilde{g}}\to G(y)\to 0$ *is a short exact sequence for all* $x\in A,y\in B$.

Proof. $F(x)=0,\forall x\in A$ since (F,A) is a trivial soft $\Gamma-$module over $\Gamma-$module M_1,so $\widetilde{f}$ is a monomorphism. $G(y)=M_2,\forall y\in B$ since (G,B) is a whole soft $\Gamma-$module over $\Gamma-$module $M_2.g:M\to M_2$ is an epimorphism as $0\to M_1\to^f\to M\to^g\to M_2\to 0$ is a short exact sequence, so $\widetilde{g}$ is an epimorphism. □

Proposition 5.3. *Let* (F,A) *be a trivial soft* $\Gamma-$*module over* $\Gamma-$*module* M_1 *and* (G,B) *be a whole soft* $\Gamma-$*module over* $\Gamma-$*module* M *if* $0\to M_1\to^f\to M\to^g\to M_2\to 0$ *is a short exact sequence, then* $0\to f(F)(x)\to^{\widetilde{f}} M\to^{\widetilde{g}} g(G)(y)\to 0$ *is a short exact sequence for all* $x\in A,y\in B$.

Proof. $F(x)=0,\forall x\in A$ since (F,A) is a trivial soft $\Gamma-$module over $\Gamma-$module M_1.Ker$f=0$, so Ker$f=F(x),\forall x\in A$,consequently $(f(F),A)$ is trivial soft $\Gamma-$module over M. (G,B) is a whole soft $\Gamma-$module over M and $g:M\to M_2$ is an epimorphism, so $(g(G),B)$ is a whole soft $\Gamma-$module over M_2, thus $0\to f(F)(x)\to^{\widetilde{f}} M\to^{\widetilde{g}} g(G)(y)\to 0$ is a short exact sequence for all $x\in A,y\in B$. □

Definition 5.3. *Let $(F, A), (G, B)$ and (H, C) are three soft $\Gamma-$modules over $\Gamma-$modules M, N and K respectively. Then we say soft $\Gamma-$ exactness at (G, B), if the following conditions are satisfied:*

i) $M \rightarrow^{f_1} N \rightarrow^{f_2} K$ *is exact,*

ii) $A \rightarrow^{g_1} B \rightarrow^{g_2} C$ *is exact,*

iii) $f_1(F(x)) = G(g_1(x))$ *for all* $x \in A$,

iv) $f_2(G(x)) = H(g_2(x))$ *for all* $x \in B$,

which is denoted by $(F, A) \rightarrow^{(f_1,g_1)} (G, B) \rightarrow^{(f_2,g_2)} (H, C)$.

In this definition, if every $(F_i, A_i), i \in I$ is soft $\Gamma-$ exact, then we say that $(F_i, A_i)_{i\in I}$ is soft $\Gamma-$ exact.

Proposition 5.4. *Let (F, A) and (G, B) are two soft $\Gamma-$modules over $\Gamma-$modules M and N respectively. If $(F, A) \rightarrow^{(f,g)} (G, B) \rightarrow 0$ is soft $\Gamma-$ exact, then (f, g) is soft $\Gamma-$ homomorphism. In particular, if $0 \rightarrow (F, A) \rightarrow^{(f,g)} (G, B) \rightarrow 0$ is soft $\Gamma-$ exact, then (f, g) is soft $\Gamma-$isomorphism.*

Proof. Since $(F, A) \rightarrow^{(f,g)} (G, B) \rightarrow 0$ is soft $\Gamma-$ exact, we have $M \rightarrow^{f} N \rightarrow 0$ and $A \rightarrow^{g} B \rightarrow 0$ are exact. Thus f and g are epimorphisms, it is clear that (f, g) is homomorphism. If $0 \rightarrow (F, A) \rightarrow^{(f,g)} (G, B) \rightarrow 0$ is soft $\Gamma-$ exact, then $0 \rightarrow M \rightarrow^{f} N \rightarrow 0$ and $0 \rightarrow A \rightarrow^{g} B \rightarrow 0$ are exact. Thus f and g are isomorphisms, it is clear that (f, g) is soft $\Gamma-$isomorphism. □

Definition 5.4. *Let $M = 0$ and $A = 0$, then $(F, A) = 0$. We call (F, A) is a zero-soft $\Gamma-$ module.*

Proposition 5.5. *Let $(F, A), (G, B)$ and (H, C) are three soft $\Gamma-$modules over $\Gamma-$modules M, N and K respectively. If $(F, A) \rightarrow^{(f_1,g_1)} (G, B) \rightarrow^{(f_2,g_2)} (H, C)$ is soft $\Gamma-$ exact with f_1, g_1 epimorphism and f_2, g_2 monomorphism, then (G, B) is a zero-soft $\Gamma-$ module.*

Proof. Since $(F, A) \rightarrow^{(f_1,g_1)} (G, B) \rightarrow^{(f_2,g_2)}\rightarrow (H, C)$ is soft $\Gamma-$ exact with f_1, g_1 epimorphism and f_2, g_2 monomorphism, we have $M \rightarrow^{f_1} N \rightarrow^{f_2} K$ and $A \rightarrow^{g_1} B \rightarrow^{g_2} C$, hence $N = 0$ and $B = 0$, it is clear that (G, B) is zero-soft $\Gamma-$ module. □

Theorem 5.1. *Let (F, A) and (H, B) are two soft $\Gamma-$modules over $\Gamma-$modules M and N respectively. For any $M \subset N, A \subset B$ and $M \subset H(x)$ where $x \in B$. If $(F, A) \rightarrow^{(f,g)} (H, B)$ is soft $\Gamma-$homomorphism, then $0 \rightarrow (F, A) \rightarrow^{(f,g)} (H, B) \rightarrow^{(f_1,g_1)} (I, B/A) \rightarrow 0$ is soft $\Gamma-$ exact, where $I(x + A) = H(x)/M$ for all $x \in B$.*

Proof. We know that $0 \to M \to^{f} N \to^{f_1} N/M \to 0$ and $0 \to A \to^{g} B \to^{g_1} B/A \to 0$ are exact. It is clear that M is a $\Gamma-$submodule of N, so that N/M is a $\Gamma-$module and M is a $\Gamma-$submodule of $H(x)$ and $H(x)/M$ is always a $\Gamma-$submodule of N/M. This shows that $(I, B/A)$ is a soft $\Gamma-$module over N/M. For all $x \in B/A$. Define $f_1 : N \to N/M$ by $f_1(n) = n + M$, for all $n \in N$. Meanwhile, we define $g_1 : B \to B/A$ by $g_1(b) = b + A$, for all $b \in B$. Therefore, it gives that

$$f_1(H(x)) = H(x) + M, I(g_1(x)) = I(x + A) = H(x) + M$$

for all $x \in B$, and hence $f_1(H(x)) = I(g_1(x))$.This implies

$$0 \to (F, A) \to^{(f,g)} (H, B) \to^{(f_1,g_1)} (I, B/A) \to 0$$

is soft $\Gamma-$ exact. □

Theorem 5.2. *Let $(F, A_2), (G, A_1)$ and (H, A) are three soft $\Gamma-$modules over $\Gamma-$modules M_2, M_1 and M respectively. If M_1 and M_2 are $\Gamma-$submodules of M with $M_2 \subset M_1, A_1$ and A_2 are $\Gamma-$submodules of A with $A_2 \subset A_1$, where $M_1 \subset H(x)$, for all $x \in A$ and $M_2 \subset G(x)$ for all $x \in A_1$. Then $0 \to (I, A_1/A_2) \to^{(f_1,g_1)} (J, A/A_1) \to^{(f_2,g_2)} (P, A/A_1) \to 0$ is soft $\Gamma-$ exact, where $I(x + A_2) = G(x)/M_2$, for all $x \in A_1$, $J(x + A_2) = H(x)/M_2$, for all $x \in A, P(x + A_1) = H(x)/M_1$, for all $x \in A$.*

Proof. Since M_1 and M_2 are $\Gamma-$submodules of M with $M_2 \subset M_1$, we have a short exact sequence $0 \to M_1/M_2 \to^{f_1} M/M_2 \to^{f_2} M/M_1 \to 0$. Since A_1 and A_2 are $\Gamma-$submodules of A with $A_2 \subset A_1$, there is a short exact sequence $0 \to A_1/A_2 \to^{g_1} A/A_2 \to^{g_2} A/A_1 \to 0$. It is clear that M_2 is a $\Gamma-$submodule of M_1, so that M_1/M_2 is a $\Gamma-$module. It gives that $G(x)/M_2$ is a $\Gamma-$module for all $x \in A_1$ from M_2 is a $\Gamma-$submodule of $G(x)$. However $G(x)/M_2$ is always a $\Gamma-$submodule of M_1/M_2. This shows that $(I, A_1/A_2)$ is a soft $\Gamma-$ module over M_1/M_2 for all $x \in A_1/A_2$. It is clear that $(J, A/A_2)$ and $(P, A/A_1)$ be a soft $\Gamma-$ module over M/M_2 and M/M_1 respectively.

Define $f_1 : M_1/M_2 \to M/M_2$ by $f_1(m_1 + M_1) = m + M_2$, for all $m_1 \in M_1$. Meanwhile, we define $g_1 : A_1/A_2 \to A/A_2$ by $g_1(a_1 + A_2) = a + A_2$, for all $a_1 \in A_1$. Therefore, we have $f_1(I(x)) = f_1(G((x)/M_2) = H(x) + M_2, J(g_1(x)) = J(x + A_2) = H(x) + M_2$ for all $x \in A_1/A_2$, so $f_1(I(x)) = J(g_1(x))$ for all $x \in A_1/A_2$.

Define $f_2 : M/M_2 \to M/M_1$ by $f_2(m + M_2) = m + M_1$, for all $m \in M$. Let $g_2 : A/A_2 \to A/A_1$ be defined by $g_2(a + A_2) = a + A_1$, for all $a \in A$. Also, we have $f_2(J(x)) = f_2(H((x)/M_2) = H(x) + M_1$ for all $x \in A/A_2$, so $f_2(J(x)) = P(g_2(x))$ for all $x \in A/A_2$. Hence $0 \to (I, A_1/A_2) \to^{(f_1,g_1)} (J, A/A_1) \to^{(f_2,g_2)} (P, A/A_1) \to 0$ is soft $\Gamma-$ exact. □

Theorem 5.3. *Let* $(F_i, A_i), i = 1,2,3,4,5$ *be a soft* $\Gamma-$*module over* $\Gamma-$*module* $M_i, i = 1,2,3,4,5$ *respectively. If* $0 \to (F_1, A_1) \to^{(f_1,g_1)} (F_2, A_2) \to^{(f_2,g_2)} (F_3, A_3) \to 0$ *and* $0 \to (F_3, A_3) \to^{(f_3,g_3)} (F_4, A_4) \to^{(f_4,g_4)} (F_5, A_5) \to 0$ *are soft* $\Gamma-$ *exact. Then* $0 \to (F_1, A_1) \to^{(f_1,g_1)} (F_2, A_2) \to^{(f_3\, f_2, g_3\, g_2)} (F_4, A_4) \to^{(f_4,g_4)} (F_5, A_5) \to 0$ *is soft* $\Gamma-$ *exact.*

Proof. Since $0 \to (F_1, A_1) \to^{(f_1,g_1)} (F_2, A_2) \to^{(f_2,g_2)} (F_3, A_3) \to 0$ and $0 \to (F_3, A_3) \to^{(f_3,g_3)} (F_4, A_4) \to^{(f_4,g_4)} (F_5, A_5) \to 0$ are soft $\Gamma-$ exact, we have $0 \to M_1 \to^{f_1} M_2 \to^{f_2} M_3 \to 0$ and $0 \to M_3 \to^{f_3} M_4 \to^{f_4} M_5 \to 0$ are exact. It is clear that $0 \to M_1 \to^{f_1} M_2 \to^{f_3\, f_2} M_4 \to^{f_4} M_5 \to 0$ is exact. Since $0 \to A_1 \to^{g_1} A_2 \to^{g_2} A_3 \to 0$ and $0 \to A_3 \to^{g_3} A_4 \to^{g_4} A_5 \to 0$ are exact. It is clear that $0 \to A_1 \to^{g_1} A_2 \to^{g_3\, g_2} A_4 \to^{g_4} A_5 \to 0$ is exact. Since $f_2(F_2(x)) = F_3(g_2(x))$ for all $x \in A_2$ and $f_3(F_3(x)) = F_4(g_3(x))$ for all $x \in A_3$. We have $f_3 f_2(F_2(x)) = f_3(F_3(g_2(x))) = F_4(g_3 g_2(x))$ for all $x \in A_2$. This implies $0 \to (F_1, A_1) \to^{(f_1,g_1)} (F_2, A_2) \to^{(f_3\, f_2, g_3\, g_2)} (F_4, A_4) \to^{(f_4,g_4)} (F_5, A_5) \to 0$ is soft $\Gamma-$ exactness. $\square$

6 Conclusion

In this work the theoretical point of view of soft $\Gamma-$ module is discussed. The work is focused on soft $\Gamma-$ module, soft $\Gamma-$ module homomorphism and soft $\Gamma-$ exactness. By using these concepts, we studied the algebraic properties of soft sets in $\Gamma-$ module structure. One could extend this work by studying other algebraic structures.

References

[1] D. Molodtsov, Soft set theory-First results, Computers Mathematics Applications, 37(4-5)(1999), 19-31.

[2] P.K. Maji, A.R. Roy, R. Biswas, An application of soft sets in a decision making problem, Computers Mathematics Applications, 44(2002), 1077-1083.

[3] D. Chen, E.C.C. Tsang, D.S. Yeung, X. Wang, The parameterization reduction of soft sets and its applications, Computer Mathematics Applications, 49(2005), 757-763.

[4] P.K. Maji, R. Biswas, A.R. Roy, Soft set theory, Computers Mathematics Applications, 45(2003), 555-562.

[5] Z. Kong, L. Gao, L. Wang and S. Li, The normal Parameter Reduction of Soft Sets and Its Algoritm, Computers and Mathematics with Applications, 56 (1) 3029-3037.

[6] J. E. Macias Diaz, On modules which are isomorphic to relatively divisible or pure submodules of each other, Quaestiones Mathematicae, 38 (6), 2015.

[7] J. Abuhlail, C. Lomp, Topological lattices and their applications to modules, Journal of Algebra Its Applications, 15 (3), (2015).

[8] R. Ameri, R. Sadeghi, Gamma modules, Ratio Mathematica 20, 127–147 (2010).

[9] A. H. Rezaei, B. Davvaz, Tensor product of gamma modules, Article in Afrika Mathematica, 26 (7), 1601- 1608, (2015).

[10] Q-M. Sun, Z-L. Zhang and J.Liu, Soft sets and modules (In Guoyin Wang, Tian-rui Li, Jerzy W. Grzymala-Busse, Duoqian Miao, Andrzej Skowron,Yiyu 54 Yao Eds.) Rough sets and knowledge technology, RSKT, Proceedings,Springer, 403-409, (2008).

[11] D.Xiang, Soft module theory, IEEE, 2013, 10th International Conference on Fuzzy Systems and Knowledge Discovery.

[12] T.Shah and S.Medhit, Primary decomposition in a soft ring and soft module, Iranian Journal of Science & Technology, 38A3 (Special issue-Mathematics) 311-320, (2014).

[13] A.Erami, A.Hasankhani and A.B.Saeid, MV-modules in view of soft set theory, Çankaya University Journal of Science and Engineering, Volume 13, No. 1 (2016) 001–015.

[14] M. I. Ali, F. Feng, X. Liu, W.K. Min, and M. Shabir, On some new operations in soft set theory, Comput. Math. Appl., 57 (9), 1547-1553, (2009).

[15] D. Pei and D. Miao, From soft sets to information systems, In proceedings of Granular Computing (Eds: X. Hu., Q. Liu, A. Skowron, T.Y.Lin, R.R. Yager, B. Zhang) IEEE, 2, 617-621.

[16] Z. Xiao, Y. Li, B. Zhong and X.Yang, Research on synthetically evaluating method for business competitive capacity based on soft set, Statical Research, 52-54.

[17] H. Aktas, N. Cağman, Soft sets and soft groups, Inform. Sci. 177 (2007), 2726- 2735.

[18] U. Acar, F. Koyuncu, B. Tanay, Soft sets and soft rings, Computers and Mathematics with Applications, 59 (11), 2010, 3458- 3463.

Ratio Mathematica **ISSN**: 1592-7415
Vol. 32, 2017, pp. 63-75 **eISSN**: 2282-8214

Teaching Least Squares in Matrix Notation

Guglielmo Monaco[1], Aniello Fedullo[2*]
[1]Department of Chemistry and Biology "A. Zambelli", University of Salerno, Italy
gmonaco@unisa.it
[2]Department of Physics "E. R. Caianiello", University of Salerno, Italy
afedullo@unisa.it

Received on: 30-05-2017. **Accepted** on: 27-06-2017. **Published** on: 30-06-2017

doi:10.23755/rm.v32i0.335

Abstract

Material for teaching least squares at the undergraduate level in matrix notation is reported. The weighted least squares equations are first derived in matrix form; equivalence with the standard results obtained by standard algebra are then given for the weighted average and the simplest linear regression. Indicators of goodness of fit are introduced and interpreted. Eventually a basic equation for resampling is derived.
Keywords: coefficient of determination, weighted sample mean, resampling, undergraduate education.
2010 AMS subject classifications: 62J05.

1 Introduction

Statistics is a never missing topic in first degree courses of scientific programs. Very soon, often at the second year undergraduate, the basic knowledge of random variables and distributions, is complemented by the simple linear regression, as a necessary tool for the interpretation of experimental data gathered in the laboratories. Indeed, the critical practice of linear regressions often forms students' basic awareness of data analysis. The advent of powerful and handy softwares on the one hand has reduced the effort required to the students for accomplishing the needed calculations, on the other hand has given them the possibility to easily perform more advanced statistical analyses [1, 2], which they cannot really understand on the grounds of the course. One of simplest of such more advanced analyses is the consideration of more regressors, the starting point of multivariate data analysis [3]. Although a specific course at the last undergraduate or first graduate year can be much profitable, we experienced that, provided the students have a basic knowledge in linear algebra, the generalized least squares can be thought at the second year undergraduate with reasonable appreciation from the class.
Reference textbooks on the matter, seemingly more diffused in the community of econometrics [5] than in that of experimental sciences [6], are not missing. However, we needed to compact some fundamental concepts and equations, and still convince the students that the more general matrix form of the least squares allows to easily retrieve the results obtainable with standard algebra. Thus, we prepared the following material, and we presented it effectively in a 12 hours module together with numerical exercises. Although our lessons obviously have a significant overlap with reference textbooks, the revised simple linear regression and the introduction of the (adjusted) weighted coefficient of determination are not easily retrieved from any of the textbooks known to us.

2 Matrix Form of the Weighted Least Squares

We consider n measures $\{y_1, y_2, ..., y_n\}$ and for each of them, say the i-th one, the regressors $\{x_{i1}, x_{i2}, ..., x_{ip}\}$, here assumed constant, which are generally coming from different associated measures. We will assume that for each measure the first regressor equals one, $x_{i1} = 1$, in order to take into account the so called intercept. The linear regression model connects the above quantities by

$$y_i \quad = \quad \sum_{j=1}^{p} x_{ij}\beta_j + \varepsilon_i \qquad i = 1, 2, ..., n \tag{1}$$

where $\beta_1, \beta_2, ..., \beta_p$ are the parameters to be estimated and $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$ are random errors, assumed independent and possibly normally distributed, with mean 0 and standard deviations $\sigma_1, \sigma_2, ..., \sigma_n$. Ordinary least squares (OLS) and weighted least squares (WLS), also called homoskedastic and heteroskedastic regressions, are the names used to distinguish the special case of equal values for all standard deviations from the case of different values. The equations for WLS of course also apply to the special OLS case.

Dividing eq. 1 by σ_i, i.e. given $z_i := \frac{y_i}{\sigma_i}$, $q_{ij} := \frac{x_{ij}}{\sigma_i}$, $\varsigma_i := \frac{\varepsilon_i}{\sigma_i}$, and using the matrix notation, the model is written as

$$z = Q\beta + \varsigma, \tag{2}$$

or, equivalently,

$$W^{\frac{1}{2}}y = W^{\frac{1}{2}}X\beta + W^{\frac{1}{2}}\epsilon,$$

where W is a diagonal matrix whose elements $W_{ii} := w_i = \sigma_i^{-2}$ are known as statistical weights, z and β are column matrices of n and p elements, respectively, Q is a matrix of dimension $n \times p$. It should be noticed that $Q\beta$ is the expectation value of z, i.e. $Q\beta =< z >$.

Under these hypotheses the least squares method gives an estimate of the model parameters by the minimization with respect to β of the functional

$$SS := \varsigma^T\varsigma = (z - Q\beta)^T(z - Q\beta) \tag{3}$$

$$= (z - Q\beta)^T(z - Q\beta) = z^Tz - 2\beta^TQ^Tz + \beta^TQ^TQ\beta, \tag{4}$$

where it has been considered that $\beta^TQ^Tz = z^TQ\beta$.

The estimates of the parameters by the least squares method are the solutions of the equations $\frac{\partial SS}{\partial \beta_i} = 0$, for $i = 1, 2, ..., p$, one for each model parameter. The computation of the derivative with respect to the vector of the parameters gives:

$$-Q^Tz + Q^TQ\beta = 0, \tag{5}$$

whose solution

$$\hat{\beta} = VQ^Tz \tag{6}$$

is, by definition, the least squares estimator of β, where $V := C^{-1}$, and $C := Q^TQ$, which we will assume always invertible.

We note that $\hat{\beta}$ is an unbiased estimator of β, indeed form eqs. 2 and 6 we have

$$< \hat{\beta} >= VQ^T < z >= VQ^TQ\beta = VC\beta = \beta. \tag{7}$$

An unbiased behavior also characterizes the weighted sample mean. Indeed, eq. 5 for $\beta = \hat{\beta}$ gives $Q^Tz = Q^T\hat{z}$ which, rewritten in the original variables, is $X^TWy = X^TW\hat{y}$. From this and from the initial hypothesis $x_{i1} = 1$, for any i, one gets $\sum_i w_i y_i = \sum_i w_i \hat{y}_i$, which divided by $\sum_i w_i$ shows that the weighted sample mean of the fitted values equals the weighted sample mean of the measures:

$$\bar{y}_w = \bar{\hat{y}}_w. \tag{8}$$

Given $\delta := \hat{\beta} - \beta$ from eqs. 6 and 7 one gets

$$\delta \; = \; VQ^T\varsigma, \tag{9}$$

which allows to easily compute the covariance matrix of the parameters, showing that it coincides with V

$$< \delta\delta^T >= VQ^T < \varsigma\varsigma^T > QV = VQ^TIQV = V,$$

where I denotes the identity matrix.

The standard deviations of the estimators of the parameters are given by the square roots of the diagonal elements of V.

Using the fitted values, one can write

$$z = \hat{z} + (z - \hat{z}) = Q\hat{\beta} + e,$$

where e is known as the vector of residuals, whose analysis is object of much concern in literature.

The fitted values are often written as

$$\hat{z} = Q\hat{\beta} = QVQ^Tz =: Hz, \tag{10}$$

where we have introduced the symmetric matrix H, which is known as hat matrix as it 'puts the hat on z'. This matrix is readily verified to be idempotent, $H^2 = QVQ^TQVQ^T = H$, a feature which readily allows to demonstrate the useful property of orthogonality of residuals and fitted values:

$$(z - \hat{z})^T\hat{z} = z^T(I - H)Hz = 0.$$

Given

$$SSE := min_\beta SS = e^T e,$$

the expansion 4, with ς in place of z and δ in place of β, can be rewritten as:

$$SSE = (\varsigma - Q\delta)^T(\varsigma - Q\delta) = \varsigma^T\varsigma - 2\delta^T Q^T\varsigma + \delta^T C\delta = \varsigma^T\varsigma - \delta^T C\delta,$$

where we have considered that $Q^T\varsigma = C\delta$ thanks to eq. 9.

Given $SSR := \delta^T C\delta$, which as SS and SSE is non-negative, the preceding equation becomes

$$SSE = SS - SSR$$

whose interpretation is that the error in the estimation of the parameters, yielding a nonzero SSR, reduces the sum of squares SS which could have been computed with the expectation value $\langle z\rangle = Q\beta$.

The average of SSE can be easily computed considering that $\delta^T C\delta = Tr\left[\delta\delta^T C\right]$, and then

$$\langle SSE\rangle \;=\; \left\langle\varsigma^T\varsigma\right\rangle - Tr\left[\left\langle\delta\delta^T\right\rangle C\right] = n - Tr\left(VC\right) = n - p,$$

known as the number of degrees of freedom, denoted by ν.

Notation. In the following $s_{xx,w}$, $S_{xx,w}$ e $s_{xy,w}$ indicate respectively the sample variance, sum of squares and weighted covariance, defined from the weighted sample mean $\bar{y}_w := \frac{\sum_i w_i y_i}{\sum_i w_i}$ in analogous manner to the corresponding unweighted means. We recall that their expressions are $s_{xx,w} = \overline{x^2_w} - \bar{x}^2_w$, $S_{xx,w} = s_{xx,w}\sum_i w_i$ e $s_{xy,w} = \overline{xy}_w - \bar{x}_w\bar{y}_w$, where $xy := (x_1 y_1, ..., x_n y_n)$.

3 Indicators for the Goodness of Fit

Besides reporting the best-fit parameters and the resulting fitted values, it is customary to give compact indicators of the goodness of fit.

A method which is widely used in the analysis of experimental data consists in the chi-squared test: the hypothesis that the model is correct is not rejected, at the appropriate level of significance, if SSE assumes values close to $\langle SSE\rangle$, i.e., for any number of parameters, if $\chi^2_r = \frac{SSE}{\nu}$ is close to 1. Values of χ^2_r larger or smaller than 1 are then considered as indicators of a poor fit or, respectively, overfitting.

A different approach considers weighted sample means. Defining the *weighted coefficient of determination* R^2_w as the square of the weighted sample correlation coefficient $\frac{s_{y\hat{y},w}}{\sqrt{s_{yy,w}s_{\hat{y}\hat{y},w}}}$ between data y and fitted values $\hat{y} = X\hat{\beta}$ and thus limited by $0 \leq R_w^2 \leq 1$, one has that

$$1 - R_w^2 = \frac{SSE}{S_{yy,w}} = \frac{s_{ee}}{s_{yy,w}}, \tag{11}$$

showing that $R_w^2 = 1$ iff $SSE = 0$, i.e. iff all residuals are zero. Therefore the greater the value of R_w^2 the better the agreement. Eq. 11 can be proven thanks to the orthogonality relation discussed above. The vector $\overline{y_w}w^{\frac{1}{2}}$, where $w^{\frac{1}{2}}$ is a column vector of elements $w_i^{\frac{1}{2}}$, is orthogonal to the vector of residuals $z - \hat{z}$, by virtue of eq. 8. Therefore the orthogonality of residuals and fitted values, eq. 2, still holds if the fitted values are translated by $\overline{y_w}w^{\frac{1}{2}}$. The vector relationship

$$z - \overline{y_w}w^{\frac{1}{2}} = (\hat{z} - \overline{y_w}w^{\frac{1}{2}}) + (z - \hat{z}), \tag{12}$$

graphically sketched in Figure 1, allows to assess that

$$S_{yy,w} = S_{\hat{y}\hat{y},w} + SSE, \tag{13}$$

whose interpretation is that $S_{\hat{y}\hat{y},w}/S_{yy,w}$ is the fraction of variability of the data explained by the knowledge of Q, i.e. by the regression, and $SSE/S_{yy,w}$ is the unexplained one, i.e. that coming from errors.

Still from eq. 12 one gets

$$S_{\hat{y}y,w} = (\hat{z} - \overline{\hat{y}_w}w^{\frac{1}{2}})^T(z - \overline{y_w}w^{\frac{1}{2}}) = (\hat{z} - \overline{\hat{y}_w}w^{\frac{1}{2}})^T(\hat{z} - \overline{\hat{y}_w}w^{\frac{1}{2}}) = S_{\hat{y}\hat{y},w} \tag{14}$$

and then

$$R_w^2 = \frac{S_{\hat{y}y,w}^2}{S_{\hat{y}\hat{y},w}S_{yy,w}} = \frac{S_{\hat{y}\hat{y},w}}{S_{yy,w}}. \tag{15}$$

Insertion of eq. 15 in eq. 13 readily gives eq. 11.

In order to discourage the introduction of models too complicated for the data examined, it has been introduced the adjusted determination coefficient

$$R_a^2 = 1 - (1 - R_w^2)\frac{n-1}{n-p},$$

obtained substituting the unbiased variances in the rhs of eq. 11.

It often happens that standard deviations of experimental data are only approximately known. A common assumption is that the standard deviations σ_i are known but for a factor k: $\sigma_i = k\tilde{\sigma}_i$, with the $\tilde{\sigma}_i$ known *a priori*. If the adjustment

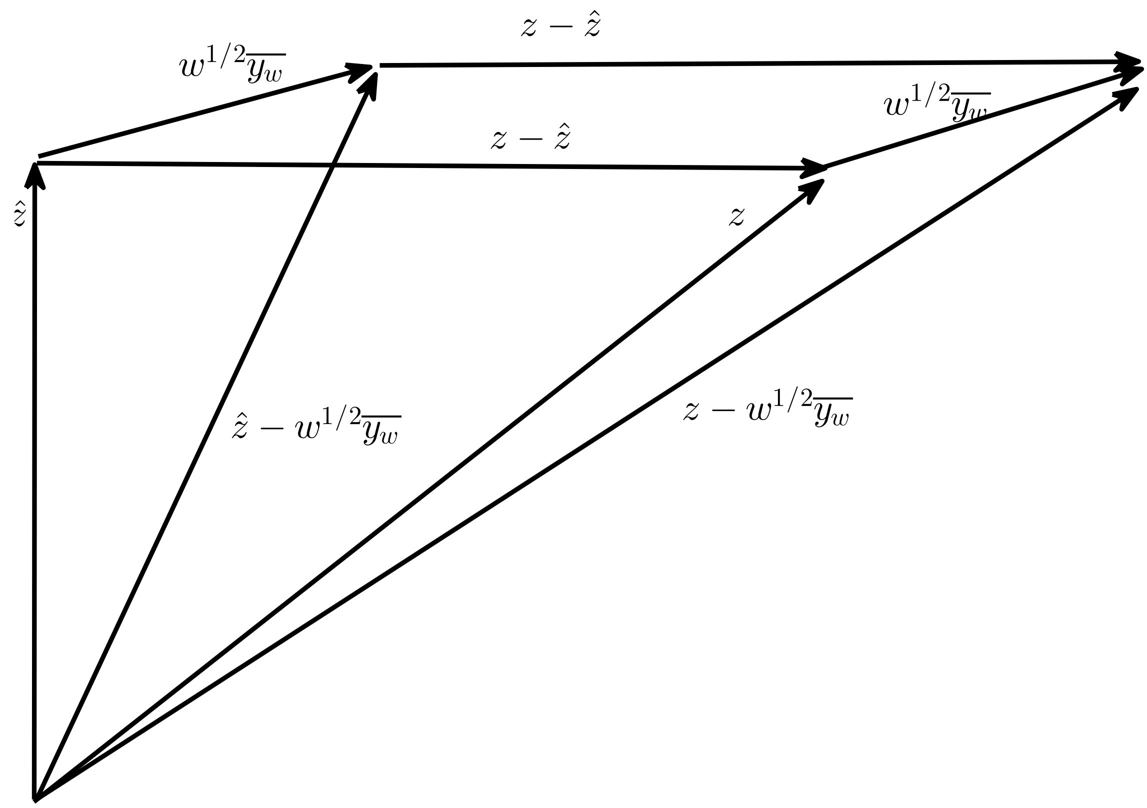

Figure 1: The residuals $z - \hat{z}$ are orthogonal to both the estimates $\hat{z}$ and the vector $\overline{y_w}w^{\frac{1}{2}}$.

of k leads to a good fitting for the model, χ_r^2 should be close to ν. Using this value, one gets

$$\nu = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{k^2 \tilde{\sigma}_i^2},$$

and a trial value for k is obtained as

$$k = \sqrt{\frac{1}{\nu} \frac{\sum (y_i - \hat{y}_i)^2}{\tilde{\sigma}_i^2}}.$$

4 Basic Applications

4.1 (Weighted) mean

The model $y = \beta \mathbf{1} + \varepsilon$ has an $n \times 1$ matrix of relative regressors, whose i-th element is

$$q_{i1} = w_i^{\frac{1}{2}}$$

Application of eq. 7 soon gives as the best fit parameter the weighted mean

$$\hat{\beta} = VQz = \frac{\sum_i w_i y_i}{\sum_i w_i} = \bar{y}_w$$

and its variance is the sum of the weights: $\sigma_\beta^2 = V_{11} = \sum_i w_i$.

4.2 WLS for a straight line

The standard linear regression considers the model $y = a + bx$. In the above notation $a = \beta_1$ and $b = \beta_2$ and the regressor matrix is

$$X = \begin{bmatrix} 1 & 1 & ... & 1 \\ x_1 & x_2 & ... & x_n \end{bmatrix}^T$$

The matrix of relative regressors will be then

$$Q = \begin{bmatrix} \sqrt{w_1} & \sqrt{w_2} & ... & \sqrt{w_n} \\ \sqrt{w_1}x_1 & \sqrt{w_2}x_2 & ... & \sqrt{w_n}x_n \end{bmatrix}^T,$$

the vector of relative data

$$z = \left[\begin{array}{cccc} \sqrt{w_1}y_1 & \sqrt{w_2}y_2 & \dots & \sqrt{w_n}y_n \end{array}\right]^T,$$

and

$$C = \sum_i w_i \left[\begin{array}{cc} 1 & \overline{x}_w \\ \bar{x}_w & \overline{x_w^2} \end{array}\right],$$

whose inverse gives the covariance matrix of the parameters

$$V = \frac{1}{S_{xx,w}} \left[\begin{array}{cc} \overline{x_w^2} & -\bar{x}_w \\ -\bar{x}_w & 1 \end{array}\right].$$

The standard deviations of the estimators of the parameters will be then

$$\left[\begin{array}{c} \sigma_{\hat{a}} \\ \sigma_{\hat{b}} \end{array}\right] = \left[\begin{array}{c} \sqrt{V_{11}} \\ \sqrt{V_{22}} \end{array}\right] = \frac{1}{\sqrt{S_{xx,w}}} \left[\begin{array}{c} \sqrt{\overline{x_w^2}} \\ 1 \end{array}\right],$$

and the estimated parameters will be

$$\left[\begin{array}{c} \hat{a} \\ \hat{b} \end{array}\right] = VQ^Tz = \frac{1}{s_{xx,w}} \left[\begin{array}{cc} \overline{x_w^2} & -\bar{x}_w \\ -\bar{x}_w & 1 \end{array}\right] \left[\begin{array}{c} y_w \\ \overline{xy}_w \end{array}\right] = \left[\begin{array}{c} \bar{y}_w - \frac{s_{xy,w}}{s_{xx,w}}\bar{x}_w \\ \frac{s_{xy,w}}{s_{xx,w}} \end{array}\right],$$

which in case of all equal weights (*homoskedastic regression*) have the simpler expression $\left[\begin{array}{cc} \bar{y} - \frac{s_{xy}}{s_{xx}}\bar{x} & \frac{s_{xy}}{s_{xx}} \end{array}\right]^T$.

4.3 Revised simple linear regression

We now give a simplified approach for the bivariate weighted linear regression: given $\mathbf{1} := [1\ 1\ \dots\ 1]^T$, we subtract $\overline{y}_w\mathbf{1}$ from the data and from the fitted data and, considering that $\overline{y}_w = \overline{\hat{y}}_w = a + b\bar{x}_w$, we obtain

$$y - \overline{y}_w\mathbf{1} = b(x - \bar{x}_w\mathbf{1}) + \varepsilon, \tag{16}$$

which, with $z := W^{\frac{1}{2}}(y - \overline{y}_w\mathbf{1})$, $q := W^{\frac{1}{2}}(x - \bar{x}_w\mathbf{1})$ e $\varsigma := W^{\frac{1}{2}}\varepsilon$, can be written as in eq. 2,

$$z = bq + \varsigma,$$

but here there is the single parameter b to be determined, as in the example of the weighted mean.

This means that matrix C is the scalar $S_{xx,w}$ readily invertible, and then $V = C^{-1} = \frac{1}{S_{xx,w}}$. On the other hand, as

$$q^T z \;=\; (x - \bar{x}_w \mathbf{1})^T W (y - \overline{y}_w \mathbf{1}) = S_{xy,w} + \overline{y}_w \sum_i w_i (x_i - \bar{x}_w) = S_{xy,w},$$

from eq. 6, one gets again $\hat{b} = \frac{s_{xy,w}}{s_{xx,w}}$. Writing now the model as $y - bx = a\mathbf{1} + \varepsilon$, the example in 4.1 gives for the intercept $\bar{y}_w - b\bar{x}_w$, from where, replacing b with its estimator[1], one finally gets $\hat{a} = \bar{y}_w - \hat{b}\bar{x}_w$, as in 4.2.

It is to be considered, however, that this simplification leads to loose information on the covariance of the a and b parameters, which should then be recover *ex post* (Appendix).

4.4 Resampling and the Best-fit Parameters

A remarkable representation of the p best-fit parameters can be obtained if one tries to determine them from the $\binom{n}{p}$ p-elements subsets of the original set of n measures [4]. Let $S_{(s)}$ be a $p \times n$ matrix obtained from the $n \times n$ identity matrix, upon selecting the p rows whose indices form subset s, with $s = 1, \dots \binom{n}{p}$. Let also $M^{[k|v]}$ be the matrix obtained from matrix M upon replacing its k-th column with vector v.

For any p-elements subset s, the data needed for the WLS are stored in vector $z_{(s)} = S_{(s)} z$ and the square matrix $Q_{(s)} = S_{(s)} Q$; the best-fit parameters are

$$\hat{\beta}_{(s)} = Q_{(s)}^{-1} z_{(s)} = X_{(s)}^{-1} W_{(s)}^{-1/2} W_{(s)}^{1/2} y_{(s)} = X_{(s)}^{-1} y_{(s)}, \tag{17}$$

which shows that, for p measures, WLS and OLS give the same results.

Use of Cramer's rule on eqs. 5 and 17 gives

$$\hat{\beta}_k = \frac{\det Q^T Q^{[k|z]}}{\det Q^T Q}, \tag{18}$$

and

$$\hat{\beta}_{(s)k} = \frac{\det Q_{(s)}^{[k|z]}}{\det Q_{(s)}} = \frac{\det X_{(s)}^{[k|y]}}{\det X_{(s)}}, \tag{19}$$

Use of the Cauchy-Binet theorem to expand the determinants of the equation 18 leads to

$$\hat{\beta}_k = \frac{\sum_s \det Q_{(s)} \det Q_{(s)}^{[k|z]}}{\sum_s \det Q_{(s)} \det Q_{(s)}} = \frac{\sum_s w_s \hat{\beta}_{(s)k}}{\sum_s w_s}, \tag{20}$$

which is the equation for a weighted average of the OLS results $\hat{\beta}_{(s)k}$ with weights

$$w_s = (\det Q_{(s)})^2. \tag{21}$$

[1]Implicit use is made of the functional invariance of the estimator $\hat{b}$.

The above representation of the best-fit parameters is the starting point for robust modifications of WLS, where the basic idea is to exclude from the mean the more extreme values of $\beta_{(s)k}$ [7].

5 Conclusions

The least squares method, a fundamental piece of knowledge for students of all scientific tracks, is often introduced considering the simple linear regression with only two parameters to be determined. However, the availability of ever more large data sets prompts even undergraduate students to a sounder and wider knowledge of linear regression. Here, we have used the linear algebra formalism to compact the main results of the least squares method, encompassing ordinary and weighted least squares, goodness of fit indicators, and eventually a basic equation of re-sampling, which could be used to stimulate interested students in an even broader knowledge of data analysis. The compactness of the equations reported above allow their introduction at the undergraduate level, provided that basic linear algebra has been previously introduced.

Acknowledgements

Financial support from the MIUR (FARB2015) is gratefully acknowledged.

References

[1] R.J. Carroll and D. Ruppert. *Transformation and weighting in regression*. Chapmand and Hall, 1988.

[2] G. Casella and R.L. Berger. *Statistical inference*. Cengage Learning, 2001.

[3] A. J. Dobson and A. G. Barnett. *An Introduction to Generalised Linear Models*. Chapman and Hall, 2008.

[4] R. W. Farebrother. Relations among subset estimators: A bibliographical note. *Technometrics*, 27(1):85–86, 1985.

[5] W.H. Greene. *Econometric Analysis*. Prentice Hall, 2012.

[6] J. R. Taylor. *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. University Science Books, 1997.

[7] C. F. J. Wu. Jackknife, bootstrap and other resampling methods in regression analysis. *The Annals of Statistics*, 14(4):1261–1295, 1986.

Appendix

Moments of $\hat{a}$ e $\hat{b}$

Averages.

$< \hat{b} >= \frac{<s_{xy,w}>}{s_{xx,w}} = \frac{(x-\bar{x}_w\mathbf{1})^T W <y-\bar{y}_w\mathbf{1}>}{s_{xx,w}} = \frac{(x-\bar{x}_w\mathbf{1})^T W <y>}{s_{xx,w}} = \frac{(x-\bar{x}_w\mathbf{1})^T W <a\mathbf{1}+bx>}{s_{xx,w}} = b\frac{(x-\bar{x}_w\mathbf{1})^T W (x-\bar{x}_w\mathbf{1})}{s_{xx,w}} = b$;

$< \hat{a} >=< \bar{y}_w > -\bar{x}_w < \hat{b} >= a + b\bar{x}_w - \bar{x}_w b = a$

The estimators are then unbiased.

Variances.

We shall use the following auxiliary results:

i) $Cov(y_i, y_j) = \delta_{ij} w_i^{-1}$

ii) $Var(\bar{y}_w) = Tr(W)^{-1}$

iii) $Cov(\bar{y}_w, \hat{b}) = 0$

Given $d := W(x - \bar{x}_w\mathbf{1})$, we have that $d^T\mathbf{1} = \sum_i d_i = 0$ and then $S_{xy,w} = d^T(y - \bar{y}_w\mathbf{1}) = d^T y$; Then $Var(S_{xy,w}) = \sum_{ij} d_i d_j Cov(y_i, y_j) = \sum_i d_i^2 w_i^{-1}$

$= \sum_i w_i (x_i - \bar{x}_w)^2 = S_{xx,w}$ from which $Var(\hat{b}) = \frac{Var(S_{xy,w})}{S_{xx,w}^2} = \frac{1}{S_{xx,w}}$.

$Var(\hat{a}) = Var(\bar{y}_w) - \bar{x}_w Cov(\bar{y}_w, \hat{b}) + \bar{x}_w^2 Var(\hat{b}) = Tr(W)^{-1} + \frac{\bar{x}_w^2}{S_{xx,w}} = \frac{\overline{x^2}_w}{S_{xx,w}}$.

.

$Cov(\hat{a}, \hat{b}) = Cov(\bar{y}_w - \bar{x}_w\hat{b}, \hat{b}) = Cov(\bar{y}_w, \hat{b}) - \bar{x}_w Var(\hat{b}) = -\frac{\bar{x}_w}{S_{xx,w}}$.

Proof of the auxiliary results

i) $Cov(y_i, y_j) = Cov(a + bx_i + \varepsilon_i, a + bx_j + \varepsilon_j) = Cov(\varepsilon_i, \varepsilon_j) = \delta_{ij}\sigma_i^2 = \delta_{ij} w_i^{-1}$.

ii) $Var(\bar{y}_w) = Tr(W)^{-2} \sum_{ij} w_i w_j Cov(y_i, y_j) = Tr(W)^{-2} \sum_i w_i = Tr(W)^{-1}$

iii) $Tr(W) Cov(\bar{y}_w, S_{xy,w}) = \sum_{ij} w_i d_j Cov(y_i, y_j) = \sum_i d_i = 0$ and then $Cov(\bar{y}_w, \hat{b}) = 0$.

$\square$

Publisher:
Accademia Piceno – Aprutina dei Velati in Teramo (A.P.A.V.)

Periodicity:
every six months

Printed in 2017 in Pescara (Italy)

Autorizzation n. 9/90 of 10/07/1990 released by Tribunale di Pescara
ISSN: 1592-7415 (printed version) -

Autorizzation n. 16 of 17/12/2013 released by Tribunale di Pescara
ISSN: 2282-8214 (online version) -

Accademia Piceno - Aprutina dei Velati in Teramo
ACCADEMIA DI SCIENZE, LETTERE, ARTI E TECNOLOGIA

www.eiris.it – www.apav.it

Ratio Mathematica, 32, 2017

Contents

Published by Accademia Piceno - Aprutina dei Velati in Teramo (A.P.A.V.)