

The Roles of Condition Number for the interval Hull in Least Squares Equation.

Stephen Ehidihamhen Uwamusi

Department of Mathematical Sciences, Kogi State University, Anyigba, Kogi State, Nigeria.

Email : Stephen_uwamusi@yahoo.com

Abstract. The paper considers nearness to singularity for which the Co-variance matrix in the least squares equation is well known. A synchronization of the condition number with ill-conditioning is highlighted which relates the quality of approximate solution to the described system. Various theoretical lower and upper bounds to the perturbed least squares problems have been described for which, reachability theory has strong representation. In particular, a theorem due to Rump as exemplified by Popova was revisited and examined in detail; a slight modification was made to the theorem by neglecting the residual iteration term appearing in the equation. As a comparison to the computed results, a procedure described in Kramer/Rohn was used to crop the corner points solution of the linear interval system from least squares equation based on the appropriate choice of the orthant (where there are 2^n possibilities), this leads to solving systems of linear inequalities for the interval Hull of solution set. Furthermore, the Rump/Krawczyk method was used to narrow, the computed corner point solution in order to obtain tighter approximate solution bounds of the interval Hull which may be applicable to both non-parametric and parametric interval linear equations. The loss function for the computed result obtained from Rump method for the set of data points is reported.

Keywords : least squares problem, diffeomorphism, reachability theory, condition number, interval arithmetic, regulator in least squares problems

MSC (2000): 65F20, 65 G40

1. Introduction

The paper considers condition numbers associated with solving least squares equations with some noise in the data. The Covariance matrix in least squares interval equations is often Ill-conditioned. Ill-conditioned problems are often encountered in geophysics, signal processing, and medical imaging ,see[1] and the cited references therein. For example, Auto-Covariance Least squares are used in the study of Kalma filter [2]. Kalma filter has been widely applied in navigation and control of vehicles, as for example, aircraft and space craft. In its widest application areas, Kalma filter is used in the navigation system of nuclear ballistic missile submarines, and cruise missiles –e.g., the US Navy’s Tomahawk missile, US Air Launched Cruise missile are a few typical examples for this purpose for which least squares equations are involved in their formulations [2]. Interval arithmetic has been found useful in the study of fractal mechanics where a fracture in any part of aircraft makes every other part inoperable. Computing reachable set is an essential aspect of nonlinear dynamic and control systems. Reachability is an essential tool in the theory of validation of interval computing, since safety is always a

paramount interest in scientific and engineering designs which takes into consideration scenarios of worst case error bounds, [3],[4],[5] and [6]

Reachability problem comes in different flavours as for example, the approximate estimates for the solution to the linear system in least squares equation may be taken as ellipsoids which depends on the operator parameter. For the exhaustive discussion on theory of reachability, the importance of topological space cannot be underscored.

The least squares problem is formulated in the form:

$$\text{Find } \beta \in R^n \ni Y = f(\beta) = \|Y - T\beta\|_2 = \min_{\beta \in R^n} \|Y - T\beta\|_2 \quad (1.1)$$

where $T \in R^{m \times n}$, $Y \in R^m$.

In what follows, we signify our notation by the convention: $R^n, IR^n, R^{m \times n}, IR^{m \times n}$ as representing real n- vector, interval n-vector, real $m \times n$ rectangular matrix and interval $m \times n$ matrix respectively. As a convention, interval numbers are represented in intervals. Thus interval T is represented as $[T, \bar{T}]$ in the form of lower and upper bounds. The midpoint of interval matrix T as well as interval vector b is respectively defined in the form

$$T_c = 0.5(T + \bar{T}),$$

$$b_c = 0.5 \begin{pmatrix} b + \bar{b} \\ - \end{pmatrix}$$

The interval radius of interval matrix [T] as well as interval vector [b] is given by

$$\Delta T = 0.5 \begin{pmatrix} \bar{T} - T \\ - \end{pmatrix}, \Delta b = 0.5 \begin{pmatrix} \bar{b} - b \\ - \end{pmatrix}$$

We note that R^n itself is convex, and all interval boxes in R^n are convex. A subset $D \subseteq R^n$ is convex if there are points $\beta, z \in D \ni t\beta + (1-t)z \in D \forall t \in [0,1]$.

Interval arithmetic operations possess strict validation of every computed result in any calculation and any such computed result is always perfectly correct when faithfully implemented. This captures all possible conversion errors, rounding errors, approximation errors which are always vigorously estimated, [6].

The separation theorem for convex set, [7], states that a point outside a closed convex set is separated from D by a hyper plane.

It is expected that readers acquaint themselves with the concept of well known implicit function theorem which gives necessary conditions under which a continuously differentiable function has an inverse. This leads us to the concept of a diffeomorphism [8], that transforms a nonlinear system in equation 1.1 to a linear system

$$T^T T \beta = T^T b \Rightarrow \hat{\beta} = (T^T T)^{-1} T^T Y = T^+ Y, \quad (1.2)$$

where $T^+ = (T^T T)^{-1} T^T$ is well defined.

It is assumed that $[X] \in ID \subset IR^m$ be closed in IR^m . A function $f : ID \rightarrow IR^n$ is a diffeomorphism of ID onto its image $Y = f(ID)$ if it is one-to-one, smooth, and of full rank n .

By full rank, would imply the maximal possible rank, e.g., the case of a rectangular matrix of size $m \times n$, the full rank is given by $\min\{m, n\}$.

Solving least squares problems in interval form is not without some combinatorial difficulties.

As a typical example, the Co-Variance matrix from system of equation 1.1 is sometimes nearly singular or, even singular.

A case of how to resolve problem of singularity in the least squares problem using the regulator parameter has been discussed [9]. The solution could be written in the form

$$\hat{\beta} = (T^T T + \tau^k I)^{-1} T^T Y, k \geq 2 \quad (1.3)$$

The term τ is taken to be $\frac{1}{\|T^+\|}$, where T^+ is the pseudo inverse operator. Interestingly, [10] proved

that the best value for k in equation 1.3 is 2.

In what follows the matrix product $T^T T$ will be denoted by another matrix B , and $T^T Y = b$ as a vector. The inverse of the matrix B is the Co-Variance matrix. It will play a major role in our analysis hence forth.

In the mean time, the paper is categorized as follows: Section 2 gives the condition numbers associated with least squares problems. Section 3 discusses the perturbation error bounds for the lower and upper chain reachable set of the approximate solution to the least squares problems wherein, the residual error plays vital roles in their formulation. In section 4 the theorem due to

[10], exemplified in [11], re-examined, modified and were adopted for a useful purpose in the realization of numerical results.

2. CONDITION NUMBERS OF LINEAR LEAST SQUARES PROBLEM

Condition number has very useful property in both Scientific and Engineering designs. It is also used in economics.

The quality of results computed from the least squares equation is based on the condition numbers of the system of equation 1.2. Condition number shows how the solution to the described system fits the data under study. Condition number measures how a change in input is propagated to change in output. High condition number implies high non linearity in the obtained solution to the system.

For a real matrix B which is not singular and whose norm $\|\cdot\|$ is known, the condition number [12,13], was defined to be

$$K(B) = \lim_{\varepsilon \rightarrow 0^+} \sup_{\|\Delta B\| \leq \varepsilon \|B\|} \frac{\|(B + \Delta B)^{-1} - B^{-1}\|}{\varepsilon \|B^{-1}\|}, \quad (2.1)$$

Whereas, the traditional condition number was known to be

$$K(B) = \|B^{-1}\| \cdot \|B\| \quad (2.2)$$

If condition number $K(B)$ of a matrix B is known in advance, it is possible to calculate the loss function, the digits of accuracy of solution in the variable β compared to precision of B by taking the logarithm of $K(B)$.

The derivation of condition number in equation (2.2) may be obtained as follows: Suppose that $\beta^{(sle,abs)}$ denotes the least squares solution to equation 1.1. Then, following [14], it is that

$$\begin{aligned} \beta^{(sle,abs)}(B) &= \lim_{\beta \rightarrow \beta_0} \sup \frac{\|\beta - \beta_0\|}{\beta^{(sle)}(x)} \\ &= \lim_{\beta \rightarrow \beta_0} \frac{\|\beta - \beta_0\| \|\beta_0\|}{\|B\beta - b\|} = \lim_{\delta\beta \rightarrow 0} \sup \frac{\|\delta\beta\| \|\beta_0\|}{\|B\delta\beta\|} = \|B^{-1}\| \cdot \|\beta_0\|. \end{aligned}$$

As a result, the relative condition number is then the quantity

$$\beta^{(sle,rel)}(B) = \frac{\|B\|}{\|\beta_0\|} \beta^{(sle,abs)}(B) = \|B^{-1}\| \|B\|.$$

The problem of computing β , given b has a condition number

$$K_\beta = \frac{\|B^{-1}\| \|\beta\|}{\|b\|}, \text{ whereas, the problem of computing b given } \beta \text{ has a condition number, } K_b = \frac{\|B\| \cdot \|\beta\|}{\|b\|}.$$

Let B be perturbed by ΔB and β be the solution to system of equation 1.1, the perturbation in b would be approximately bounded by

$$\frac{\|\delta b\|/\|b\|}{\|\Delta B\|/\|B\|} \approx K_b \leq K(B) \quad (2.3a)$$

$$\|\delta b\| \leq \frac{K(B)\|b\| \cdot \|\Delta B\|}{\|B\|} \quad (2.3b)$$

Similarly, assuming B is perturbed by ΔB and b is given, then the perturbation in β is also approximately equal to :

$$\frac{\|\delta \beta\|/\|\beta\|}{\|\Delta B\|/\|B\|} = K_\beta \leq K(B) \Rightarrow \|\delta \beta\| \approx \frac{K(B)\|\beta\|\|\Delta B\|}{\|B\|} \quad (2.4)$$

The product norm for the data space which holds for $\|\beta\|_2$ as the solution space, [15] is given by the equation

$$\|(B, b)\|_F = \sqrt{\|B\|_F^2 + \|b\|_2^2} \quad (2.5)$$

Note that $\|B\|_F$ represents Frobenius norm and is defined as $\|A\|_F = \left[\sum_{i=1}^m \sum_{j=1}^n |b_{ij}|^2 \right]^{\frac{1}{2}}$ while the P-norm

$\|B\|_P = \sup_{\beta \neq 0} \frac{\|B\beta\|_P}{\|\beta\|_P}$ for $P = 1, 2, \infty$. We note in passing that both the Frobenius and P-norms satisfy

certain inequalities in the form:

$$\|B\|_2 \leq \|B\|_F \leq \sqrt{n} \|B\|_2, \quad (2.6)$$

$$\text{Max } |b_{ij}| \leq \|B\|_2 \leq \sqrt{mn} \max |b_{ij}|, \quad (2.7)$$

$$\|B\|_2 \leq \sqrt{\|B\|_1 \cdot \|B\|_\infty} \quad (2.8)$$

Via the definition of relative condition number, as the quantity

$$\beta^{(sl,abs)}(B) = \frac{\|B\|}{\|\beta_0\|} \beta^{(sl,abs)}(B) = \|B^{-1}\| \|B\|, \quad \text{the relative condition number obtained in [16,17] using}$$

some iterative methods is given by the quantity

$$K^{(rel)}(B, b) = \|B^+\|_2 \left(\|B^+\|_2^2 \|r\|_2^2 + \|\beta\|_2^2 + 1 \right)^{\frac{1}{2}} \frac{\|(B, b)\|_2}{\|\beta\|_2}, \quad (2.9)$$

for any $r = b - B\beta$.

The reciprocal of condition number is equal to the distance to the nearest singular matrix for all structured perturbations.

Most importantly, [13] used the Bauer-Skeel condition number for a weighted matrix to show that

$$Cond_{BS}(B, E) = \left\| |B^{-1}| E \right\| \quad (2.10)$$

E, being defined as $|B|$ and relates condition number with spectral radius by the equation:

$$Cond(B, E) = \rho(|B^{-1}| E) = D_1^{-1} D_2 \inf_{D_1, D_2} \rho(D_1 B D_2) \quad (2.11)$$

D_1, D_2 are diagonal scaling and ρ is the spectral radius. It follows [12] that

$$\frac{1}{\rho(|B^{-1}| E)} \leq \delta(B, E) < \frac{(3 + 2\sqrt{2})}{\rho(|B^{-1}| E)} \quad (2.12)$$

The expression $\delta(B, E)$ in equation 2.12 is the familiar singular value of (B, E) . In any case, [12] warned that the factor $(3 + 2\sqrt{2})$ cannot be replaced by 1 since computationally, $\delta(|B|, E) = \frac{n}{\rho(|B^{-1}| |B|)}$.

3. Backward perturbation error bound for Least squares equation

The smallest Frobenius perturbation ΔB which makes a given non-zero β into a solution space of the perturbed problem $\min_{\beta} \|b - (B + \Delta B)\beta\|_2$ in a sense due to [16] was estimated to be

$$X_F^{(LS)}(\beta) = \left(\frac{\|r\|_2^2}{\|\beta\|_2^2} + \min\{0, \lambda\} \right)^{\frac{1}{2}} \quad (3.1)$$

Where ,

$$\lambda = \lambda_{\min} \left(BB^T - \frac{rr^T}{\|\beta\|_2^2} \right) \quad (3.2)$$

The number λ is the smallest eigenvalue of the $m \times n$ matrix. The lower bound for equation 3.1 was found to be

$$\frac{1}{\sqrt{2}} X_F^{(LS)}(\beta) \leq X_2^{(LS)} \leq X_F^{(LS)}(\beta) \quad (3.3)$$

More complicated upper error bound for the least squares solution set was later found in [18] to be

$$\left(X_2^{(LS)}(\beta)\right)^2 = W \frac{(r^T B\beta)^2}{\|\beta\|_2^2 (\|B\beta\|_2^2 + \|r\|_2^2)} \quad (3.4)$$

Wherefrom,

$$W = \frac{2}{1 + \sqrt{1 - \eta}}, \quad (3.5)$$

$$\eta = \frac{4(r^T B\beta)^2}{(\|B\beta\|_2^2 + \|r\|_2^2)^2} \quad (3.6)$$

Two conditions may be adduced for least squares problem to be ill-conditioned namely, either b could be nearly orthogonal to the column space of B , or, B itself is ill-conditioned, [19].

In what follows, given that $BB^+ \Delta B = \Delta B$, and $B^+ B(\Delta B)^T = (\Delta B)^T$, if $\text{range}((\Delta B)^T) \subseteq \text{range}(B^T)$ and $\text{range}(\Delta B) \subseteq \text{range}(B)$, then $(B + \Delta B)^T$ can be approximated [20] by the truncated Taylor Series as

$$(B + \Delta B)^T = B^+ - B^+ (\Delta B) B^+ + O(\varepsilon^2) \quad (3.7)$$

Therefore, since

$$\beta + \delta\beta = (B + \Delta B)^+ (b + \delta b) = \beta + B^+ \delta b - B^+ (\Delta B) B^+ b + O(\varepsilon^2), \quad (3.8)$$

$$\delta b = -\frac{\|b\|}{\|B\|\|\beta\|} \Delta B \beta, \quad (3.9)$$

$\delta\beta$ be approximated by the equation

$$\delta\beta = -B^+ \Delta B \beta \left(1 + \frac{\|b\|}{\|B\|\|\beta\|} \right) \quad (3.10)$$

Thus the condition number for the solution set is

$$K^S(B, b) = \|B^+\| \|B\| + \|B^+\| \cdot \frac{\|b\|}{\|\beta\|} \approx \|B^+\| \|B\| + \|B^+\| \frac{\|B\beta\|}{\|\beta\|} \quad (3.11)$$

4. Computability of Reachable Sets

As said earlier at the beginning of the paper, computability set is a very important issue in nonlinear dynamic and control theory, [21,22]. Reach-ability problem comes in different flavours, for instance, given a diffeomorphism which transforms a nonlinear system to a linear system, the approximate estimates for the solution set may be considered in the form of ellipsoids which depends on the operator. Various values of parameter will lead to a family of ellipsoids whose intersection is the exact reach set based on some good iterative methods.

As it were, one can apply ellipsoid on linear system to obtain the reach set where the inverse operator becomes very necessary to approximate the nonlinear system for the convex bodies as in calibration of stereo, [22].

In most cases, the only worry one has is “ which is the best operator” as a tool to achieve this objective? Thus we need the best selection of interval Hull $\sum (B, b)$ as solution set to the given problem 1.1.

In what follows, after transforming problem 1.1 into equivalent problem 1.2, we introduce the parametrised linear interval system in the sense of [10 and 11] given by:

$$B(p).\beta = b(p) \quad (4.1)$$

where $B(p) \in R^{m \times n}$ and $b(p) \in R^n$ all depend on parameter vector $p \in R^k$, and p varies within a range $[p] \in IR^k$, the set of solution to all $B(p).\beta = b(p), p \in [p]$. The interval hull to the parametrised linear interval system 4.1 is described in the sense of [10] which is expressed in the form:

$$\sum .^p = \sum (B(p), b(p), [p]) = \{\beta \in R^n \mid B(p).\beta = b(p), (for\ some\ p \in [p])\}. \quad (4.2)$$

Where, each individual component $\{B(p)\}_{ij}$ and $\{b(p)\}_j$ of $B(p)$, respectively, $b(p)$ depends linearly on p means that there are vectors

$$\omega(i, j) \in R^k \text{ for } 0 \leq i \leq n, 1 \leq j \leq \text{with}$$

$$\{B(p)\}_{ij} = \omega(i, j)^T . p \text{ and } \{b(p)\}_j = \omega(0, j)^T . p$$

Such occurrences are often encountered in engineering practices, [11].

In what follows, we review the following theorem for an invariant set such that, $F(\beta) \subset \beta$. This relates Brouwer Fixed point theorem to contraction mapping. We are interested in an interval vector β that is chain reachable from β_0 to β for which the $\mathcal{E}$ -chain holds good for problem 1.1.

The chain reachable set F from β_0 is defined as $\text{Chain-Reach}(F, \beta_0) = \bigcap_u \text{Reach}(F, \beta_0, u)$ where u runs over all open covers of β when iterative methods are applied.

Theorem 4.1,[11]. Let $B(p).\beta = b(p)$ with $B(p) \in R^{n \times n}, b(p) \in R^n, p \in R^k$ be a parametrised linear system, where $B(p), b(p)$ are given. Let $R \in R^{n \times n}, [Y] \in IR^n, \beta \in R^n$, define $[Z] \in IR^n, [C] \in IR^{n \times n}$ by

$$[Z]_i = \left(\sum \left\{ R_{ij} \cdot \left(\omega(0, j) - \bar{\beta}_v \cdot \omega(j, v) \right) \right\} \right)^T [p],$$

$$[C] = I - R.B([p]).$$

Define $[V] \in IR^n$ by means of the following iteration enclosure

$$1 \leq i \leq n : V_i = ([Z]) + [C].[U],$$

$$[U] = (V_1, V_2, \dots, V_{i-1}, Y_i, \dots, Y_n)',$$

If, $[V] \overset{\subset}{\neq} [Y]$, then R and every matrix $B(p), p \in [p]$ are regular, and for every $p \in [p]$ the unique solution

$$\hat{\beta} = B^{-1}(p).b(p), \quad (4.3)$$

satisfies

$$\hat{\beta} = \beta_c + [V] \quad (4.4)$$

Considering equation 4.4, we make a slight change to rewrite the Rump/Krawczyk method that will be used in the form:

$$[\hat{\beta}] = Rb_c + (I - R[B])([\beta] - \beta_c) \quad (4.5)$$

In method 4.5, the term R appearing is a pre-conditioner of a matrix $[B]$. The reason being that method 4.5 is more suggestive to employ than method described in theorem 4.1

Following discussions in [23], we crop the corner point solution of $\sum ([B], [b]) \bigcap O_k$, where O is the sign vector, $s = (s_i) \in S^n$, $(s_i \in \{-1, +1\}, i=1(1)n$, the corresponding signs of the components of an interior point of O , O is an orthant, [23,24]. In [23], the exact solution to the linear interval system of

equation 1.2 obtained from least squares problem 1.1 could be obtained by solving half spaces computable $\bar{H}_i, \bar{H}_i$ systems of linear inequalities which are dependent on the choice of orthant O (2^n possibilities), results of which can be further sharpened by method 4.5 and is represented in the form:

$$\sum ([B], [b]) \cap O = \bigcap_{i=1}^n \left(H_i \cap \bar{H}_i \right) \cap O. \quad (4.5)$$

In [25], Alefeld and Mayer denoting $\bar{H}_i, \bar{H}_i$ by $H_i^k, \bar{H}_i^k, k=1(1)2^n$ gave representation for exact solution set of the linear interval system of equation 4.5 in the form:

$$\sum ([B], [b]) = \bigcup_{k=1}^{2^n} \left[\bigcap_{i=1}^n \left(H_i^k \cap \bar{H}_i^k \right) \cap O_k \right] \quad (4.6)$$

The expression in the right hand side bracket is either $\emptyset$ or a convex polytope. As a result, we are led to the following theorem an extension of Oettli-Prager theorem [4,6] which narrows closely what is discussed in equations 4.5 and 4.6.

Theorem 4.2, [26]. Let $[B] = [B_c - \Delta B, B_c + \Delta B]$ be an $n \times n$ interval matrix, $[b] = [b_c - \delta, b_c + \delta]$ an interval n -vector, where ΔB and δ are as defined in section 1. Let Z be a subset of S_n having the following properties:

(i) $\text{sgn}(\beta_0) \in Z$ for some $\beta_0 \in \sum(B, b)$,

(ii) for each $z \in Z$ the inequalities

$$(QB_c - I)O_z \geq |Q|\Delta B, \quad (4.7)$$

$$(QB_c - I)O_{-z} \geq |Q|\Delta B, \quad (4.8)$$

have matrix solutions Q_z and Q_{-z} , respectively.

(iii) if $z \in Z$, $Q_{-z}b_c - |Q_{-z}|\delta \leq Q_z b_c + |Q_z|\delta$, and $(Q_{-z}b_c - |Q_{-z}|\delta)_j (Q_z b_c + |Q_z|\delta)_j \leq 0$

for some j , then $z - 2z_j e_j \in Z$.

Then $[B]$ is regular and

$$\sum ([B], [b]) \subseteq \bigcup_{z \in Z_0} [\beta_z, \bar{\beta}_z] \quad (4.9)$$

holds, where

$$Z_0 = \left\{ z \in Z \mid \sum ([B], [b]) \cap R_z^n \neq \emptyset \right\} \quad (4.10)$$

As a follow up to this discussion, it could be seen that theorem 4.2 coincides with the interval Hull of equation 4.6 when viewed from Oettli-Prager theorem, [26].

Let us note that an improvement on the computed results from methods 4.5 and 4.10 can be made by further using Rump/Krawczyk method [23]

$$Rb_c + (I - R[B])z^k \subset \text{int}(z^k), (z^k \supset \sum ([B], b)). \quad (4.11)$$

One iteration of method 4.11 as a refining stage would suffice to obtain an accurate approximate solution.

5. NUMERICAL EXAMPLE

Planetary elliptical orbit of ten observations of its position in the (X,Y) plane is considered based on the sample data taken from [27].

TABLE 1

X	Y
1.02	0.39
0.95	0.32
0.87	0.27
0.77	0.22
0.67	0.18
0.56	0.15
0.44	0.13
0.30	0.12
0.16	0.13
0.01	0.15

Take interval uncertainty as data noise for both dependent variable Y_i and independent variable X_i as $[\epsilon] = [-0.0005, 0.0005]$, where $i=1,2,\dots,m$. We use a polynomial of order three for the data fit.

Using cubic polynomial fit, the result for the Least squares problem 1 without noise in the data is

$$\hat{\beta} = (0.1550, -0.2578, 0.4694, -0.0295)^T.$$

Results computed from the Table 1 using the Theorem due to [10] are hereby displayed in Table 2

TABLE 2 SHOWING RESULTS COMPUTED FROM TABLE 1.

RESULT FROM POPOVA METHOD [10] FOR $\hat{\beta}$	
[0.1549994,	0.15500067]
[-0.25780655,	-0.2577921]
[0.46940375,	0.46940565]
[-0.02950075,	-0.0294949]

TABLE 3 SHOWING RESULTS COMPUTED FROM METHOD 4.5 AND IMPROVED RESULTS WITH RUMP/KRAWCZYK METHOD

CORNER POINT SOLUTION WITH CROPPING FROM METHOD 4.5 FOR $\hat{\beta}$	IMPROVED RESULTS WITH RUMP/ KRAWCZYK [10] $\hat{\beta}$
[0.1553, 0.1553]	[0.1552891, 0.1553112]
[-0.2611, -0.2620]	[-0.2617332, -0.2614201]
[0.4772, 0.4803]	[0.4787567, 0.4790004]
[-0.0367, -0.0342]	[-0.0355532, -0.0353243]

As a consequence, the value of $\frac{\|\beta_c - \hat{\beta}\|}{\|\hat{\beta}\|}$ was computed and the loss function for method [10] was

found to be [6.8446, 6.8446].

The condition number for the solution without noise using equation 3.11 was found to be 1.1684×10^4 .

The polynomial fit of order three used for [10] is given in the form:

$$f(\beta) = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$$

$$= [0.1549994, 0.15500067] + [-0.25780655, -0.2577921]X + [0.46940275, 0.46940565]X^2 + [-0.02950075, -0.0294949]X^3$$

To demonstrate the faithfulness of our implementation to the described method, we take as an example, arbitrary $X = 0.12$ and computed $f(\hat{\beta}) = [0.1308, 0.1308]$ for method [10]. This coincides within the values of the vector Y from the given data set. The same analogy goes for Table 3 as given in equation 4.5. All results were computed using MATLAB Version 2007.

6. CONCLUSION

The paper considered the roles of condition number in the studies of least squares interval equations using polynomial fit of order three. In the body of the paper, Possible areas of applications of least squares were highlighted as extension to the backward perturbation errors in least squares problems also discussed. A diffeomorphism is made to relate the reachability theory for the computable set of resulting linear interval system from which, an ellipsoid may be applied. Here, in the paper, a section of the theorem due to [10], as exemplified by Popova in [11], was adopted which provides validated bounds to the linear interval system obtained from the Least squares equation. The loss function was computed based on the approximate results from Table 2.

As can be expected, the computed results in Table 2 approximate sufficiently close to the theoretical floating point results which were obtained from the data in the absence of noise in the data using MATLAB 2007. In other words, at the computation stage when the data in Table 1 were afflicted with uncertainty, with our described method whose results were presented in Table 2, had all computed results with guaranteed error bound accounted for which narrowed closely the idea poised in [6].

Besides the above mentioned facts in the computation of results with guaranteed error bounds, we also took into consideration theoretical approximation technique for cropping corner points solution in the sense of [23,24,25] in the paper to advance the approximate solution. This involves solving systems of inequalities where there are 2^n orthants. The Rump/Krawczyk method was used to narrow further the computed approximate corner point solution method as showed in Table 3.

It is strongly suggested in our findings that same results would be obtained if theorem 4.2, attributed to [26] were used instead. Further insight is given on how to estimate $f(\hat{\beta})$ based on the given value of input data in order to further ascertain the computed approximate solution from the described numerical method.

REFERENCES

- [1] Uwamusi,SE., and Uwamusi A.E (2013). The Core error problem in variables for statistical least squares with interval data. British Journal of Mathematics and Computer Science, 4(2):226-240.
- [2] Kalmer filter-Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/kalmer_filter.
- [3] Uwamusi, SE., (2008). On the interval hull of solution sets of parameterized nonlinear equations. Scientific Research and Essay Vol. 3(9), pp.183-189. Available online at <http://www.academicjournals.org/SRE>
- [4] Uwamusi,SE., (2009). Verified bounds for nonlinear systems via Hansen-Sengupta Method, International Journal of Physical Sciences, 4(10): 571-576 , <http://www.academicjournals.org/ijps>, In INTLAB ref, www.ti3.tu-harburg.de/rump/intlab/INTLABref.pdf).
- [5] Uwamusi SE., (2013). Hansen-Bliek-Rohn method applied on Total Least Squares with interval data input, The International Journal of Engineering and Science (IJES), 2(10): 103-108. Issn(e):2319-1813, ISSN(p):2319-1805
- [6] Rump SM (2003). Algorithm for computing validated results (In J. Grabmaier, E. Kaltofen and V. Weisofennig, eds.), Computer Alg. Hand book, Chapter 2, 12.2, pp 110-112, Springer.
- [7] Neumaier A (1990). Interval methods for systems of equations, Cambridge University Press,Cambridge.
- [8] Kass RE and Vos PW (1997). Diffeomorphisms and the Inverse Function Theorem, Geometric Foundations of Asymptotic Inference, John Wiley and Sons ,Inc
- [9] Ishteva M, Usevich K and Markovsky I (2013).Regularized Structured low-rank approximation with applications, arXiv:1308-1827V1 [Math.NA] 8 August.
- [10] Rump MS (1994). Verification Methods for Dense and Sparse Systems of Equations, in Topics in Validated Computations, J. Herzberger, ed., Elsevier Science B.V , PP. 63-135.
- [11] Popova ED (2001).On the solution of parametrised linear systems. Scientific Computing, Validated Numerics, Interval Methods, (in eds Wolff von Gudenberg), Kluwer/Plenum Publishers, New York, pp. 127-138.
- [12] Rump SM (1999). Ill-conditioned matrices are componentwise near to singularity, SIAM Review 41(1): 102-112.

- [13] Rump SM (1997). Theorems of Perron-Frobenius type for matrices without sign restrictions. *Linear Algebra and its Applications (LAA)*, 266 :1-42.
- [14] Grcar JF (2003). Optimal Sensitivity Analysis of Linear Least Squares. Lawrence Berkeley National Laboratory report LBNL-52434
- [15] Arioli M, Baboulin M and Gratton S (2007). A partial Condition Number for Linear Least Squares Problems, *SIAM J. Matrix Anal, Appl.*, 29(2): 413-433.
- [16] Gu M (1999). New fast algorithms for structured linear least squares problems, *SIAM J. on Matrix Analysis and Applications*, 20(1):244-269.
- [17] Gu M (1999). Backward perturbation bounds for linear least squares problems, *SIAM J. on Matrix Analysis and Applications*, 20(2):363-372.
- [18] Walden B, Karlson R and Sun JG (1995). Optimal backward perturbation bounds for the linear least squares problems, *Numerical linear Algebra with Applications*, 2(3):271-286.
- [19] Stewart GW (1977). On the perturbation of Pseudo-inverses, projection and linear least squares problems, *SIAM Review*, 19(4):634-662.
- [20] Xu W, Wei Y and Qiao S (2006). Condition numbers for structured least squares problems, *BIT Numerical Mathematics*, DOI: 10.1007/S 10543-006-0049-0
- [21] Collins P (2005). Continuity and computability of reachable sets. *Theoretical Computer Science* 341, 162-195.
- [22] Collins P and Goldsztejn A (2008). The reach and evolve algorithm for reachability analysis of nonlinear dynamical systems, *Electronic Notes in Theoretical Computer Science* 223, 87-102.
- [23] Kramer W (2006). Computing and Visualizing solution sets of interval linear systems, Preprint BUW-WRSWT 2006/8, Wissenschaftliches Rechnen/Softwaretechnologie Fachbereich C (Mathematik und Naturwissenschaften), Bergische Universität Wuppertal, Germany.
- [24] Rohn J (1989). Systems of Linear Interval Equations. *Linear Algebra and Its Application* 126, 39-78.
- [25] Alefeld G, Mayer M (1995). On the symmetric and unsymmetric solution set of interval systems. *SIAM J. Matrix Anal. Appl.*, Vol. 16, 1223-1240.
- [26] Rohn J (2010). A General method for enclosing solutions of interval linear equations. Institute of Computer Science, Academy of Sciences of the Czech Republic, Technical report No. V-1067.
- [27] Heath MT (1997). *Scientific Computing: An Introductory Survey*, McGraw-Hill, New York.

