Ilorin Journal of Science
Volume 4, Number 2, 2017, pp. 284 – 297 (Printed in Nigeria)
ISSN: 2408 – 4840 © 2017 Faculty of Physical Sciences, University of Ilorin
https://doi.org/10.54908/iljs.2017.04.02.007

ILORIN
JOURNAL OF SCEINCE

ILJS-17-030

Solving Rectangular Nonlinear System of Equation by Filtered Tikhonov Regularization

Uwamusi*, S. E.

Department of Mathematics, University of Benin, Benin City, Edo State, Nigeria.

Abstract

The paper presents methods for solving a large-scale rectangular system of equations based on Tikhonov regularization procedures wherein, incorporated, the Singular Value Decomposition (SVD) as basis of numerical computation. We obtain the regularization parameter by adopting the Penrose-pseudo-inverse process with a view to diminishing occurrence of huge condition number appearing in the left-hand side of the equation for meaningful solution. We obtain the rank of a rectangular matrix as well as approximation of Low rank matrix, a very important tool in image reconstruction from the noisy data. It is demonstrated that the symmetric matrix coming from the normal equation is reduced to a tridiagonal matrix using the Givens- QR - transformation process wherein, the norm of the inverse tridiagonal matrix may be obtained in an economical way

Keywords: Least squares problem, Tikhonov regularization method, Singular Value Decomposition (SVD), Low rank matrix approximation, Inverse of a tridiagonal matrix

1. Introduction

One important problem in mathematics as well as in other engineering practices is fitting data to a mathematical model or finding equilibrium point to a given chemical reactions problem. This often leads to overdetermined system of nonlinear equations

$$F(x) = 0\,, \qquad (1)$$

where $F : D \subset R^m \to R^n$ is a differentiable function on $S = \bar{S}\left(x^{(0)}, r\right) \subset D$ which is assumed to be Lipschitz continuous on D, and $m > n$. Basic method for solving this is the Gauss-Newton –Tikhonov operator approximated by a linear model in the form:

$$\hat{F}(x) = F(x_k) + F^{/}(x_k)(x - x_k)\,. \qquad (2)$$

Starting with Newton method, we obtain solution to equation (1) in the form:

$$x^{(k+1)} = x^{(k)} - (F^{/}(x^{(k)})^T F^{/}(x^{(k)}))^{-1} F^{/}(x^{(k)})^T F(x^{(k)})\,.$$

*Corresponding Author: Uwamusi, S. E.
Email: stephen_uwamusi@yahoo.com

For the typical un-regularized least squares problem case, the solution is in the form

$$x^{(k+1)} = x^{(k)} - F^{/}(x^{(k)})^{+} F(x^{(k)}), \qquad (k = 0,1,..,.). \tag{3}$$

The $F^{/}(x^{(k)})^{+}$ so defined is the well-known Pseudo-inverse matrix.

By further setting $A(x^{(k)}) : S \subset R^n \to L(R^m, R^n)$, $\bar{F}(x^{(k)}) = F^{/}\left(x^{(k)}\right)^T F\left(x^{(k)}\right)$, then equation (1) can be rewritten in the form:

$$x^{(k+1)} = x^{(k)} - (A(x^{(k)})^T A(x^{(k)}))^{-1} \bar{F}\left(x^{(k)}\right), \qquad (k = 0,1,2,...,.). \tag{4}$$

For the general case, it is extended to include the square nonlinear system when $m = n$ .We then impose, a stringent condition in the sense of Ortega and Rheinboldt (2000) on Newton method:

$$x^{(k+1)} = x^{(k)} - A(x^{(k)})^{-1} F(x^{(k)}), k = 0,1,...,. \tag{5}$$

Assuming $F : D \subset R^n \to R^n$ is F is differentiable on a convex set $D_0 \subset D$ and that $\| F^{/}(x) - F^{/}(y) \| \le \eta \|x - y\|, \forall x, y \in D_0$. Let there exists a $x^{(0)} \in D_0$ such that $\|F^{/}(x^{(0)})^{-1}\| \le \beta$ and $\alpha = \beta\eta\varsigma \le \frac{1}{2}$ where $\varsigma \ge \|F^{/}(x^{(0)})^{-1} F(x^{(0)})\|$. If $P(t) = \frac{1}{2}\beta\varsigma t^2 - t + \varsigma = 0$ is the quadratic equation with roots $t_1 = (\beta\varsigma)^{-1}\left[1 - (1 - 2\alpha)^{\frac{1}{2}}\right]$, $t_2 = (\beta\varsigma)^{-1}\left[1 + (1 - 2\alpha)^{\frac{1}{2}}\right]$. It follows that for $\bar{S}\left(x^{(0)}, t_1\right) \subset D_0$, the iterates of Newton method are well defined, remain in $\bar{S}\left(x^{(0)}, t_1\right)$ and converge to a solution x^* of $F(x) = 0$ that is unique in $S\left(x^{(0)}, t_2\right) \cap D_0$.

As is well known, Xu and Chang (1997), solving a nonlinear system of equation (1) is synonymous to finding the global minimizer of $f(x) = \left(\frac{1}{2}\right) F(x)^T F(x)$. In any case, the condition that $F(x)$ be locally Lipschitz continuous implies that $f(x)$ does. The remaining section in the paper is arranged as follows: Section 2 deals with the iterated filtered Tikhonov regularization method. We made a synchronization of SVD with the Filtered Tikhonov regularization method for the resulting least squares problem. In section 3, a bound is constructed for the Singular values of an overdetermined matrix in the sense of Rump (2012) which has special importance to the described method for which a negative log likelihood function for Tikhonov regularization parameter becomes handy as a useful tool. We also discussed in the paper the Givens-QR factorization as an alternative method to checking the accuracy of approximate solution as described in section2. Section 4 gives the numerical illustration with the described procedures. In section 5, we conclude the paper based on the strength of findings.

2. The Iterated Filtered Tikhonov Regularization Method

The singular value decomposition of a matrix $A \in R^{m\times n}$ for $m > n$ in line with Uwamusi (2017) is given by $A = U\sum V^T = \sum_{j=1} u_j \sigma_j v_j^T$ where $U = (u_1, u_2, ..., u_n), V = (v_1, v_2, ..., v_n)$ are matrices with orthonormal columns, $U^T U = I_n$, $V^T V = I_n$, and $\sum = diag(\sigma_1, \sigma_2, ..., \sigma_n)$ has a non-increasing order such that $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \geq 0$ with condition number of $A = \frac{\sigma_1}{\sigma_n}$ becomes severe for $\sigma_n \to 0$. We give special features of the matrix A in the form:

$AA^T = (U\sum^T V^T)(V\sum U^T) = U(\sum^T \sum)U^T$, and $A^T A = (V\sum^T U^T)(U\sum V^T) = V(\sum^T \sum)V^T$,

$(AA^T)u_i = \sigma_i^2 u_i$, $(A^T A)v_i = \sigma_i^2 v_i$.

In order to simplify details, it is that: U is the matrix of eigenvectors for $A^T A$ while V is the matrix of eigenvectors for AA^T.Furthermore, matrices AA^T and $A^T A$ have the same positive eigenvalues. The singular values σ_j of A will cluster at zero with resultant huge condition number see e.g., Bjorck (2009), Rump (2012), because in terms of SVD

$$x_j = \sum_j^n \frac{\xi_j}{\sigma_j} v_j, \qquad (\xi_j = u_j^T b) \tag{6}$$

has coefficients ξ_j decreasing faster than σ_j.

The Tikhonov regularization method is used to solve highly ill-posed system of equations. It is used to stabilize an unstable linear system that is highly singular whose solution by traditional methods is meaningless due to presence of noisy data in the right-hand side of linear ill-posed system. The reason for this is that, there is a clustered eigenvalue close to zeros, hence the need for regularization. Choosing a best regularization parameter is a nontrivial problem. The best-known Tikhonov regularization type method is the Filtered Tikhonov regularization that is dependent on Singular values of the matrix.

The Tikhonov regularization and the side constraint are defined in the form:

$$x_\tau = \arg\min \left\{ \|Ax - b\|_2^2 + \tau^2 \|L(x - x^*)\|_2^2 \right\} . \tag{7}$$

The scalar τ is a regularization parameter that is designed to control the weight assigned to the minimization of residual norm while the L may be tken as an Identity matrix. The solution to the normal equation for least squares problem in the sense of equation (7) is then given in the form:

$$x_{\lambda} = (A^T A + \tau^2 I)^{-1} A^T b. \tag{8}$$

We present the filter function Chung *et al*. (2012, 2015), Bjorck (2009), Erhel *et al*. (2001) for Tikhonov regularization as follows:

$$x_{tik}(\phi_A) = \sum_{j=1}^{k} \left(\phi(A) \frac{u_j^T b}{\sigma_j} v_j \right) = V diag(U^t b) \Sigma^{-1} \phi_A. \tag{9}$$

The $\phi(A)$ appearing in equation (9) is given by the equation:

$$\phi(A) = \frac{\sigma_j^2}{(\sigma_j^2 + \tau^2)} . \tag{10}$$

In matrix form, we write the filter function for Truncated Singular values decomposition (TSVD) for Tikhonov regularization method assuming $\sum_r = (\sigma_1, \sigma_2, ..., \sigma_r \ )$ in the form:

$$x_{tikh}(\phi_A) = V \begin{pmatrix} \sigma_1^{-1} \frac{\sigma_1^2}{(\sigma_1^2 + \tau^2)} & 0... & & 0 \\ 0 & \sigma_2^{-1} \frac{\sigma_2^2}{(\sigma_2^2 + \tau^2)} & 0... & 0 \\ ... & ... & & ... \\ & & . & \\ & & & . \\ & & & . \\ 0 & ... & & \sigma_k^{-1} \frac{\sigma_k^2}{(\sigma_k^2 + \tau^2)} \end{pmatrix} U^T b \ . \tag{11}$$

The interpretation of equation (11) to equation (6) is that, for a very small number τ , $\phi((A))$ is approximately $\phi(\frac{1}{\sigma_j}); j = 1,2,..., k$ and, this will be noisy for small enough σ_j. On the other hand, a direct approach to regularized method of Equation (7) using SVD of $A = U \sum V^T$ is

$$(V \Sigma^2 V^T + \tau^2 I)x = V \Sigma U^T b, \tag{12}$$

with solution given by

$$x_{\lambda} = V(\Sigma^2 + \tau^2 I)^{-1} \Sigma U^T b. \tag{13}$$

In the implementation of solution to Equation (13) using MATLAB, we should take into cognizance that: $\sum = (\sigma_1, \sigma_2, ..., \sigma_n)^T$, $p = \left(U^T b\right)./\sum$, $\varphi_i = \frac{\sigma_i^2}{\left(\sigma_i^2 + \tau^2\right)}$, so that $x_\lambda = V(\varphi.*p)$.

3. Construction bound for Singular values of a Matrix A.

Fundamental to this discussion is the approach due to Rump (2012). There, it was established that "singular value decomposition $A = U\sum V^T$ reveals most important properties of A, from condition number over the distance to singularity to the solution of a linear or, in case of a rectangular matrix, both under/over determined or least squares problem''.

Theorem, Rump (2011). Let $A \in R^{m\times n}$ be given, and suppose $\left\|I - A^T A\right\| \le \alpha < 1$. Then for $m \succ n, A$ has full rank, and

$$\sqrt{1-\alpha} \le \sigma_i(A) \le \sqrt{1+\alpha} \text{ with } \frac{1}{\sqrt{1+\alpha}} \le \sigma_i(A^+) \le \frac{1}{\sqrt{1-\alpha}}$$

for all $1 \le i \le n$. In particular:

$$\sqrt{1-\alpha} \le \|A\| \le \sqrt{1+\alpha} \text{ and } \frac{1}{\sqrt{1+\alpha}} \le \|A\| \le \frac{1}{\sqrt{1-\alpha}}.$$

Now consider the model

$$Ax = b + \varepsilon, \tag{14}$$

which has $\varepsilon \approx N(0, \sigma^2 I)$ and whose variance σ^2 is unknown. Solving linear least squares problem would lead to inversion of highly ill-conditioned matrix which pushes the data noise to the right-hand side, thereby rendering solution process in most cases useless for any meaningful uses. By setting $A(\tau)$ as $A(\tau) = A\left(A^T A + \tau^2 I\right)^{-1} A^T$, the negative log likelihood function is defined by the equation

$$\Im\left(x^{reg}(\tau^2), A, b\right) = -\log \prod_{i=1}^{n} \left(\frac{1}{\hat{\sigma}\sqrt{2\pi}} \exp\left(\frac{-\left(Ax^{reg}(\tau^2) - b\right)_i^2}{2\hat{\sigma}^2} \right) \right)$$

$$= \frac{m}{2}\log 2\pi \hat{\sigma}^2 + \frac{\left\|Ax^{reg}(\tau^2) - b\right\|^2}{2\hat{\sigma}^2} \quad . \tag{15}$$

Thus when $\hat{\sigma}^2$ is the maximum likelihood estimate to $\frac{\left\|Ax^{reg}(\tau^2) - b\right\|^2}{n}$, then

$$\Im\left(x^{reg}(\tau^2), A, b\right) = \frac{n}{2}\log\frac{2\pi}{n} + \frac{n}{2}\log\left\|Ax^{reg}(\tau^2) - b\right\|^2 + \frac{n}{2}. \tag{16}$$

The bias correction term is given by $B(\tau^2, A, b) = p^{eff}(\tau) = TrA(\tau^2)$. The matrix $A^T A$ is symmetric and positive semi-definite, and by the ideas expressed in Neumaier (1998) the matrix $A^T A + \tau^2 I$ has its eigenvalues in the interval $\left(\tau^2, \tau^2 + \|A\|^2\right)$ with condition number less than or equal to $\left(\frac{1}{\tau^2}\left(\tau^2 + \|A\|^2\right)\right)$ which decreases as τ increases.

The estimate for the condition number $K(A)$ is given by the inequality

$$K(A) \le \sqrt{\left(\sigma_1^2 + \tau^2\right)}\left(\frac{\tau}{(\sigma_n^2 + \tau^2)} + \max_{1\le i\le n}\frac{\sigma_i}{(\sigma_i^2 + \tau^2)}\right). \tag{17}$$

In the QR Factorization method without regularization parameter, the matrix A is decomposed into the form $A = Q\begin{bmatrix} R \\ O \end{bmatrix}$, so that $\begin{bmatrix} R \\ O \end{bmatrix}x = Q^T b$. Since $Q^T Q = I$ and the column of Q are orthonormal. It follows Q preserves the Euclidean norm. The formulation of least squares problem with the aid of QR assumes the form

$$\|r\|_2^2 = \|b - Ax\|_2^2 = \left\|b - Q\begin{bmatrix} R \\ O \end{bmatrix}x\right\|_2^2 = \left\|Q^T b - Q^T Q\begin{bmatrix} R \\ O \end{bmatrix}x\right\|_2^2 = \left\|Q^T b - \begin{bmatrix} R \\ O \end{bmatrix}x\right\|_2^2,$$

by orthogonality of Q, implies that $Q^T Q = I$. We now give error estimates for the condition number to the perturbed system:

$$(A + \Delta A)x = b + \Delta b, \tag{18}$$

where ΔA and Δb respectively are the perturbation to A and b and are expressed in the form:

$$\eta(\hat{x}) = \min\left\{\varepsilon \,\middle|\, (A + \Delta A)\hat{x} = b + \Delta b\right\}. \tag{19}$$

Using the fact that $\|\Delta A\| \le \varepsilon\|A\|, \|\Delta b\| \le \varepsilon\|b\|$, we then compute the ratio

$$\eta(\hat{x}) = \frac{\|r\|}{\|A\|\left\|\hat{x}\right\| + \|\Delta b\|}, \quad (r = b - A\hat{x}). \tag{20}$$

In what follows we choose z to be a vector such that $z^T\hat{x} = \left\|\hat{x}\right\|\|z\| = 1$ and z is a dual of $\hat{x}$. The optimal perturbation for ΔA and Δb respectively are given by

$$\Delta A = \frac{\|A\|\left\|\hat{x}\right\|}{\|A\|\left\|\hat{x}\right\| + \|b\|} r z^T \quad , \quad \Delta b = -\frac{\|b\|}{\|A\|\left\|\hat{x}\right\| + \|b\|} r \, .$$

The backward error in the sense of Walden *et al*. (1995) for the perturbed problem of Equation (18) is described in the form:

$$\eta\left(\hat{x}\right) = \min\left\{\|\Delta A, \theta\Delta b\|_F \,\middle|\, (A + \Delta A)\hat{x} - (b + \Delta b)_2 = \min! \right. \quad . \tag{21}$$

Provided that:

$$\eta\left(\hat{x}\right) = \left(\frac{\|r\|_2^2}{\left\|\hat{x}\right\|}\mu + \min\{0, \lambda_*\}\right)^{\frac{1}{2}} \text{, and } \lambda_* = \lambda_{\min}\left\{AA^T - \mu\frac{rr^T}{\|x\|_2^2}\right\}. \tag{22}$$

$$\mu = \frac{\theta^2\|x\|_2^2}{1 + \theta^2\left\|\hat{x}\right\|_2^2} \text{, and } \|(\Delta A, \theta\Delta b)\| = \sqrt{\|\Delta A\|_F^2 + \|\Delta b\|_2^2} \, .$$

Particularly, is the case when the real eigenvalue $\lambda < 0$. Then we would have that

$$\left(\frac{\|r\|_2^2}{\left\|\hat{x}\right\|}\mu + \lambda_*\right)^{\frac{1}{2}} = \delta_{\min}([A.R])$$

and

$$R = u^{\frac{1}{2}}\frac{\|r\|_2}{\left\|\hat{x}\right\|_2}\left(I - rr^T\right) \, ,$$

$$\theta = \frac{\|A\|_F}{\|b\|_2}, \; x^+ = \frac{x^T}{(x^T x)}.$$

4. Numerical Examples.

Problem 4.1:

We consider the rectangular matrix $A \in R^{m \times n}$ where $m > n$ defined by

$$A=\begin{pmatrix}
1 & 68 & 4624 & 314432 & 21381376 \\
1 & 68 & 4624 & 314432 & 21381376 \\
1 & 74 & 5476 & 405224 & 29986576 \\
1 & 83 & 6889 & 571787 & 47458321 \\
1 & 83 & 6889 & 571787 & 47458321 \\
1 & 83 & 6889 & 571787 & 47458321 \\
1 & 48 & 2304 & 110592 & 5308416 \\
1 & 37 & 1369 & 50653 & 1874161 \\
1 & 51 & 2601 & 132651 & 6765201 \\
1 & 61 & 3721 & 226981 & 13845841 \\
1 & 64 & 4096 & 262144 & 16777216 \\
1 & 56 & 3136 & 175616 & 9834496 \\
1 & 33 & 1089 & 35937 & 1185901 \\
1 & 41 & 1681 & 68921 & 2825761 \\
1 & 48 & 2304 & 110592 & 5308416 \\
1 & 84 & 7056 & 592704 & 49787136 \\
1 & 80 & 6400 & 512000 & 40960000 \\
1 & 82 & 6724 & 551368 & 45212176 \\
1 & 78 & 6084 & 474552 & 37015056 \\
1 & 79 & 6241 & 493039 & 38950081 \\
1 & 60 & 3600 & 216000 & 12960000 \\
1 & 55 & 3025 & 166375 & 9150625 \\
1 & 65 & 4225 & 274625 & 17850625 \\
1 & 73 & 5329 & 389017 & 28398241 \\
1 & 75 & 5625 & 421875 & 31640625 \\
1 & 76 & 5776 & 438976 & 33362176 \\
1 & 67 & 4489 & 300763 & 20151121 \\
1 & 76 & 5776 & 438976 & 33362176 \\
1 & 80 & 6400 & 512000 & 40960000 \\
1 & 80 & 6400 & 512000 & 40960000 \\
1 & 81 & 6561 & 531441 & 43046721 \\
1 & 82 & 6724 & 551368 & 45212176 \\
1 & 82 & 6724 & 551368 & 45212176 \\
1 & 84 & 7056 & 592704 & 49787136
\end{pmatrix}$$

$$b=(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)^T.$$

Knowing well that the rank approximation to A in the spectral norm is $\|.\|_2$. We set that $A_r = \sum_{i=1}^{r} \sigma_i u_i v_i^T$, to have that $A_r = \min_{A_r} \|A - A_r\|_2 = \left\| \sum_{i=r+1}^{n} \sigma_i u_i v_i^T \right\|_2 = \sigma_{k+1}$.

Similarly, we have that

$$\|A - A_r\|_F^2 = \left\| \sum_{i=r+1}^{n} \sigma_i u_i v_i^T \right\|_F^2 = \sum_{i=r+1}^{n} \sigma_i^2 . \tag{23}$$

The question is "can we estimate the upper bound for the Pseudo inverse matrix A^+"? To do this will require the use of SVD of A and using relevant ideas due to Bjorck (2009), for example. Because $A = U \sum V^T$, $\sum = diag(\sigma_1, \sigma_2, ..., \sigma_n)$, inductive argument implies $A^T A + \tau^2 I = V\left(\sum^2 + \tau^2 I\right)V^T$. The Pseudo inverse matrix $A^+ = (A^T A)^{-1} A^T$ is transformed in a manner analogous to the form:

$$\left(A^T A + \tau^2 I\right)^{-1} A^T = V\left(\sum^2 + \tau^2 I\right)^{-1} \sum U^T . \tag{24}$$

If we take the norm of both sides of Equation (24), then we see that

$$\left\| \left(A^T A + \tau^2 I\right)^{-1} A^T \right\| = \max_i \frac{\sigma_i}{\sigma_n^2 + \tau^2} . \tag{25}$$

The other pertinent details as a fall out to Equation (25) are:

$$\left\| \left(A^T A + \tau^2 I\right)^{-1} \right\| = \frac{1}{\sigma_n^2 + \tau^2} , \tag{26}$$

$$\|A\|_2 = \max_x \frac{\|Ax\|}{\|x\|} = \left(\max_x \frac{\|Ax\|^2}{\|x\|^2} + \tau \right)^{\frac{1}{2}} = \left(\sigma_1(A)^2 + \tau^2\right)^{\frac{1}{2}} . \tag{27}$$

Turning back to the Numerical problem 4.1 above, the following results are obtained. The computed results are displayed in Table 1 below, where we used the Penrose pseudo-inverse process as Tikhonov regularization parameter in our solution.

Table 1: Showing Results for the linear system.

Results from Normal Equation QR method $\hat{x}$	**Results From Regularized Tikhonov parameter** $\tau = 0.0182$ x_{tik}	**Results from using SVD** $\hat{x}$
$\begin{pmatrix} 1.0000 \\ 0.0000 \\ 0.0000 \\ 0.0000 \\ 0.0000 \end{pmatrix}$	$\begin{pmatrix} 0.0531 \\ 0.0691 \\ -0.0018 \\ 0.0000 \\ -0.0000 \end{pmatrix}$	$\begin{pmatrix} 1.0000 \\ -0.0000 \\ 0.0000 \\ -0.0000 \\ 0.0000 \end{pmatrix}$

The matrix A has huge condition number $K(A) = \sigma_1 / \sigma_n = 1.9391e + 11$, with singular values

$$\Sigma = 1.0e + 08 * \begin{pmatrix} 2.0748 & 0 & 0 & 0 & 0 \\ 0 & 0.0107 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \end{pmatrix},$$

where the orthonormal matrix $V = \begin{pmatrix} -0.0000 & 0.0000 & 0.0007 & 0.0535 & -0.9986 \\ -0.0000 & 0.0002 & 0.0305 & 0.9981 & 0.0535 \\ -0.0001 & 0.0125 & 0.9994 & -0.0308 & -0.0009 \\ -0.0102 & 0.9999 & -0.0125 & 0.0002 & 0.0000 \\ -0.9999 & -0.0102 & -0.0000 & -0.0000 & -0.0000 \end{pmatrix}$,

for the ill-conditioning of linear system of problem 4.1.

Particularly, we also, showed the result for the Low rank matrix approximation for Problem 4.1 in the form:

$$\min_{\hat{z}} \|ZA - I_n\|^2 = (1.0e + 09) * \begin{pmatrix} 0.0000 & 0.0000 & 0.0000 & -0.0011 & -1.7725 \\ -0.0000 & -0.0000 & 0.0000 & 0.0001 & 0.1059 \\ 0.0000 & 0.0000 & -0.0000 & -0.0000 & -0.0019 \\ -0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ -0.0000 & -0.0000 & -0.0000 & -0.0000 & -0.0000 \end{pmatrix}$$

In the course of implementation, we noted Xiang and Zou (2013),that the condition number, $K(A)$ associated to the Tikhonov regulator parameter defined earlier has the distinct quality:

$$K(A) = \|A\|\|A^{+}\| \leq \sqrt{\sigma_i^2 + \tau^2}\left(\frac{\tau}{\sigma_n^2 + \tau^2} + \max_{1 \leq i \leq n} \frac{\sigma_i}{\sigma_i^2 + \tau^2}\right) . \quad (28)$$

We make special remarks on the occurring matrix $B = A^T A$. For purposes of analysis, Hargreaves (2006), using givens orthogonal matrix plane rotations, Uwamusi and Otunta (2002) assuming the matrix B is dense, it is possible to reduce this matrix B to tridiagonal in the form:

$$T = \begin{pmatrix} \alpha_0 & \beta_1 & & & & & & \\ \beta_1 & \alpha_1 & \beta_2 & & & & & \\ & \beta_2 & \alpha_2 & \beta_3 & & & & \\ & & . & . & . & & & \\ & & & . & . & . & & \\ & & & & . & . & . & \\ & & & & & & \alpha_{n-1} & \beta_{n-1} \\ & & & & & & \beta_{n-1} & \alpha_n \end{pmatrix} . \quad (29)$$

The inverse of this matrix T is a useful tool for analysis in numerical analysis. To compute $\|T^{-1}\|_1$ in $O(n)$ operations, the following approach is adopted: After sometime, the transformation for the matrix R is equal to the form:

$$R = \begin{pmatrix} r_1 & h_1 & t_1 & & & & & \\ & r_2 & h_2 & t_2 & & & & \\ & & . & . & . & & & \\ & & & . & . & . & & \\ & & & & . & . & . & \\ & & & & & r_{n-2} & h_{n-2} & t_{n-2} \\ & & & & & & r_{n-1} & h_{n-1} & t_{n-1} \\ & & & & & & & & r_n \end{pmatrix} \quad (30)$$

while the matrix Q is equal to

$$Q^T = \begin{pmatrix} i_1 j_1 & \sin\theta_1 & & 0 & \\ i_2 j_2 & i_2 j_2 & \sin\theta_2 & & \\ . & . & . & & \\ . & . & . & & \\ . & . & . & & \\ i_n j_1 & i_n j_2 & \ldots & \ldots i_n j_{n-1} & i_n j_n \end{pmatrix} . \quad (31)$$

For clarity of purposes, it is that:

$$D = diag\left(\left(1, -\sin\theta_1, \sin\theta_1 \sin\theta_2 \ldots (-1)^{n-1}\right)(\sin\theta_1 \sin\theta_2 \ldots \sin\theta_{n-1})\right),$$

$$i = D^{-1}(1, \cos\theta_1, \cos\theta_2, \ldots, \cos\theta_{n-1})^T, \; j = D(\cos\theta_1, \cos\theta_2, \ldots, \cos\theta_{n-1}, 1)^T .$$

5. Conclusion

The paper presented solution for solving a large-scale rectangular system of equation. Particularly important, is the choice of Tikhonov regularization parameter using the Penrose pseudo-inverse process. Solving ill-conditioned system of equations is often fraught with problem of obtaining meaningless solution because of huge condition number being pushed to the right-hand side of the system. This paper discussed processes for overcoming this problem. The SVD is numerically backward stable; hence its applications in many areas have been presented. For example, we used this, to obtain the rank of a rectangular matrix and approximation of Low rank matrix, a very important tool in image reconstruction from the noisy data. The upper bound for the condition number of the matrix A was discussed using ideas due to Rump (2012). Sample numerical problem has been demonstrated in section 4. All computed results are displayed in Table 1. Thus, from Table 1, the computed results from QR and SVD for normal equation are the same. We used Tikhonov regularized method with SVD to obtain result in column 2 in Table 1. Therefore, it follows that the obtained results for Regularized Tikhonov method gave more meaningful estimate in a reasonable sense.

References

Bjorck, A. (2009): *Numerical methods in Scientific Computing*. Volume II, SIAM.

Chung, J., Chung, M. and O'Leary, D. (2012): Optimal Filters from Calibration data for image deconvolution with data acquisition error. *Journal of Mathematical Imaging and Vision* **44** (3), 366-374.

Chung, J., Chung, M. and O'Leary, D. (2015): Optimal regularized low rank inverse approximation. *Linear Algebra and its applications*. **468**, 260-269.

Erhel, J., Fédéric, G. F. and Saad, Y. (2001): Least squares polynomial filters for ill-conditioned linear systems. Report No. 4175 Theme 4, Institut National De recherché en informatiqueet en Automatique, France.

Hargreaves, G. (2006): Computing the condition number of tridiagonal and diagonal-plus-semiseparable matrices in linear time. *SIAM Journal on Matrix Analysis and Applications.* **27** (3), 801-820.

Neumaier, A. (1998): Solving Ill-conditioned and singular linear systems: A Tutorial on regularization. *SIAM Review*. **40**, 636-666.

Ortega, J. M. and Rheinboldt, W. C. (2000): *Iterative Solution of Nonlinear equations in several variables. Classics in Applied Mathematics*. SIAM Philadelphia, USA.

Rump, S. M. (2011): Verified bounds for singular values, in particular for the spectral norm of a matrix and its inverse. *BIT Numerical Mathematics*, **51** (2), 367-384.

Rump, S. M. (2012): Verified bounds for least squares problems and underdetermined linear Systems. *SIAM Journal on Matrix Analysis and Applications* **33** (1), 130-148.

Uwamusi, S. E. and Otunta, F. O. (2002): Computation of Eigenvalues of Hermitian matrix Via Givens Plane Rotation. *Nigerian Journal of Applied Sciences*. **20**, 101-106.

Uwamusi, S. E. (2017): On denoising solution space to least squares problems. *Transactions of the Nigerian Association of Mathematical Physics*. **5**, 73-78.

Walden, B., Karlson, R. and Sun, J. G. (1995): Optimal backward perturbation bounds for the linear least squares problem. *Numerical Linear Algebra with Applications.* **2** (3), 271-286.

Xiang, H. and Zou, J. (2013): Regularization with randomised SVD for large scale Inverse problems. *Inverse Problems*. **29**, 085008.

Xu, H. and Chang, X. W. (1997): Approximate Newton methods for nonsmooth equations. *Journal of Optimization Theory and Application*. **93** (2), 373-394.