International Journal of Mathematical Archive-4(7), 2013, 119-132

REGULARIZATION OF ILL-CONDITIONED LINEAR SYSTEMS

Joseph Acquah & Ernest Danso-Addo

Department Of Mathematics, University Of Mines and Technology, Tarkwa

(Received on: 15-05-13; Revised & Accepted on: 03-07-13)

ABSTRACT

In this paper, we examined various regularization methods for obtaining meaningfully approximations to ill-conditioned linear systems. The methods considered include; Preconditioning, Truncated Singular Value Decomposition (T.S.V.D) and Tikhonov Regularization. These methods were applied to obtain stable solution estimates for the Hilbert system, Boundary Value Problems (B.V.P) and Discretized Fredholm Integral Equation of the First Kind. Also, different methods for choosing an optimal regularization parameter including the L-Curve method were considered.

Keywords: Ill-Conditioned, Preconditioning, Truncated Singular Value Decomposition, Tikhonov Regularization and L-Curve.

1. INTRODUCTION

The linear system problem Ax = B arises in many branches of applied mathematics, computational sciences and engineering. [1]. The most common source of these problems is in the numerical solutions of ordinary and partial differential equations, as well as integral equations [2]. Often, if the matrix A is inherently ill-conditioned or the righthand side B is contaminated with noisy data, due to measurement error or by a slight perturbation, the computed solution is usually a meaningless approximation to the exact solution. The need arises to incorporate some information about the desired solution in order to stabilize the problem to obtain a useful and more stable solution, [6]. That is, we replace the original problem with a nearby problem to reduce the effects of perturbation error or noise-free data in the solution. One way of doing that, is by constraining the solution to minimize the error in the solution to obtain a desired solution. The process of incorporating this additional information in order to obtain a meaningful solution to the problem is called regularization [3]. The concept of regularization has been looked at by different researchers.

Some of the works addressing the problem from a different angle are the work of Arnold [7] and Hansen [8]. In this article, we compare three methods for stabilizing the linear system

$$Ax = B \tag{1}$$

The methods considered are preconditioning, Truncated singular value decomposition and Tikhonov regularization method.

First, we solve system (1) by preconditioning the matrix A and the vector B by a pre-conditioned matrix M, to improve the convergence of the solution. The system is then examined with an example.

In the second case, we consider minimizing the error in the system (1) by truncating the small singular values σ_i that magnify round-off errors, of the least square solution

$$\hat{x}_{ls} = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i$$
⁽²⁾

of equation (1).

Finally, we apply the Tikhonov regularization method, coupled with the L-Curve to compute the exact or least squares solution of linear in (1) by minimizing the quadratic function

$$\phi_{\lambda}(x) = \|Ax - B\|_{2}^{2} + \lambda \|Lx\|_{2}^{2}, \qquad (3)$$

Corresponding author: Adetunde Isaac

International Journal of Mathematical Archive- 4(7), July – 2013

Joseph Acquah & Ernest Danso-Addo / REGULARIZATION OF ILL-CONDITIONED LINEAR SYSTEMS/ IJMA- 4(7), July-2013.

subject to the side constraint, to reduce the error in the solution. The matrix L is an identity, a diagonal weighting matrix or a (n-p)x n discrete approximation of the p^{th} derivative operator and λ is a regularization parameter, greater than zero. Our computational results shows that the later generates meaningful solution approximate than the first two approaches.

2. CONDITIONING OF A PROBLEM

Conditioning refers to the sensitivity of the solution of any giving problem to small changes in the input data, [5].

Let P(x) denote the value of a problem corresponding to input data x and δx denotes a small perturbation in x, then P is said to be **ill-conditioned**, if the relative error in the solution is much larger than the relative error in the data. That is:

$$\frac{\left|P(x+\delta x)-P(x)\right|}{\left|P(x)\right|} \gg \frac{\left|\delta x\right|}{\left|x\right|}$$
(4)

For many problems, a **condition number** can be defined. If the condition number is *large*, then the problem is said to be **ill-conditioned**. On the other hand, if the condition number is *small*, then the problem is said to be wellconditioned.

Consider the problem of computing a function y = f(x).

Suppose we are interested in the effects on $y \in Y$ when a given $x \in X$ is perturbed slightly by a *small* amount δx ,

then the relative size of the perturbation in x is $\frac{|\delta x|}{|x|}$, and it's corresponding relative size of the perturbation in f(x)

can be written as

$$\frac{\left|f(x+\delta x)-f(x)\right|}{f(x)}\approx\frac{\left|\delta xf'(x)\right|}{\left|f(x)\right|}=\frac{\left|xf'(x)\right|}{\left|f(x)\right|}\times\frac{\left|\delta x\right|}{\left|x\right|}$$

The quantity

$$k = \frac{\left|xf'(x)\right|}{\left|f(x)\right|} \tag{5}$$

is called the condition number for the problem. If the quantity is large, the problem is ill-conditioned; on the other hand, if it is small, the problem is well-conditioned.

Condition Number of a Matrix

Consider the special case where f in equation (5) is a linear function as in (1), then the k can be estimated as follows:

$$k = \lim_{\Delta x \to 0} \left[\frac{\left\| A(x + \Delta x) - Ax \right\|}{\left\| Ax \right\|} / \frac{\left\| \Delta x \right\|}{\left\| x \right\|} \right]$$

This can be simplify as:

$$k = \lim_{\delta x \to 0} \left[\frac{\|A(x + \delta x) - Ax\|}{\|Ax\|} / \frac{\|\delta x\|}{\|x\|} \right] = \|A\| \frac{\|x\|}{\|Ax\|}.$$
(6)

But

$$||x|| = ||Ix|| = ||AA^{-1}x|| = ||A^{-1}Ax|| \le ||A^{-1}||||Ax||$$

Or further

$$\frac{\left\|x\right\|}{\left\|Ax\right\|} \le \left\|A^{-1}\right\|. \tag{7}$$

Combining equation (6) and (7) we obtain $k \le \|A\| \|A^{-1}\|$.

But for certain value of
$$\alpha$$
, it can be deduced that
 $k = \alpha \|A\| \|A^{-1}\|,$
(9)

where α is proportionality constant.

3. THE HILBERT SYSTEM AND ITS ACCURACY

The Hilbert system is an n x n square matrix denoted by H_n and composed of fractional entries with the largest values located for small values of I and j, with entries defined as

$$H_n = h_{ij} = \frac{1}{i+j-1}$$
, where $1 \le i \le n$, and $1 \le j \le n$, (10)

called the Hilbert Matrix of order n.

The Hilbert matrix arises in least squares polynomial approximation of continuous functions on the interval [0, 1], using the standard basis $1, x, x^2, \dots, x^n$ for P^n . Thus, the matrix $H_n = \lfloor h_{ij} \rfloor$ is easily identified as the Hilbert matrix since

$$h_{ij} = \int_{0}^{1} x^{i+j-2} dx = \frac{1}{i+j-1}$$
(11)

4. A FORMULATED PROBLEM

Consider the problem of solving the linear system Hx = B, where H is a 12 x 12 Hilbert matrix, and B chosen such that the linear system has the exact solution x = [1, 1, 1, 1, 1, 1, 1, 1, 1, 1].

If the system is solved, using any of the standard method such as the Exact Inverse, QR, LU, and Cholesky factorizations, we realized that for small values of n, the computed solution (\hat{x}) is reasonably accurate. However as n increases the precision degenerates very rapidly.

	$\ x - \hat{x}\ _{\infty}$		No. of Digits Lost
n	$\ x\ _{\infty}$	Cond(A)	
5	$7.135919719e^{-12}$	$4.766072502e^{+05}$	4
6	$1.540944388e^{-11}$	$1.495105864e^{+07}$	5
7	$3.016719939e^{-10}$	$4.753673562e^{+08}$	6
8	$4.623435736e^{-08}$	$1.525757542e^{+10}$	8
9	$5.40907181e^{-06}$	$4.931538214e^{+11}$	10
10	$1.80367381e^{-04}$	$1.602515828e^{+12}$	12
11	$6.28747466e^{-03}$	$5.221040338e^{+14}$	13
12	$7.42237815e^{-01}$	$1.794510255e^{+15}$	15

Table 1: Accuracy of a Hilbert System.

Table 1 shows the error, the condition number and the number of digits lost for selected values of n. We realize from the Table 1, that the exact solution of the system is accurate to about 16 significant digits. For small values of n, the

(8)

computed solution shows a significant digit of accuracy. When $n \ge 12$, the computed solution (\hat{x}), has little or no significant digits of accuracy.

In effect, we can deduced that a matrix H has condition number of order 10^{K} if H and B are accurate to d significant digits. The computed solution \hat{x} to Hx = B is also said to be accurate to (d - k) digits. If k is greater than d, then the computed solution has no digits of precision.

5. REGULARIZATION METHODS FOR LINEAR ILL-POSED PROBLEMS

Regularization methods are used to obtain meaningful solution estimates for discrete ill-posed problems or rankdeficient linear problems [9]. In cases where some parameters are ill-determined either by least-square methods or in situations where the number of parameter is larger than the number of available measurements, it is necessary to stabilize the system by using regularization methods. Some of these methods are: Truncated Singular Value Decomposition (TSVD), Preconditioning and Tikhonov Regularization.

6. TRUNCATED SINGULAR VALUE DECOMPOSITION

The idea behind truncated singular value decomposition is to replace all the small singular values less than certain threshold say, α with exact zeros. This reduces the error in the solution that contaminates the solution and also reduces the order of the matrix. Thus in theory, the singular value decomposition (S.V.D) of a matrix A, is a decomposition of the form

$$\mathbf{A} = U \sum V^{T} = \sum_{i=1}^{n} U_{i} \sum_{i} V_{i}^{T}$$

If A is invertible, then its inverse is given by

$$\mathbf{A} = \sum_{i=1}^{n} v_i \boldsymbol{\sigma}_i^{-1} \boldsymbol{u}_i^T,$$

where

$$U = [u_1, ..., u_n], V = [v_1, ..., v_n], and \Sigma = diag[\sigma_i, ..., \sigma_n].$$

Therefore, the solution to Ax = B can be defined as

$$x = \sum_{i}^{n} \sigma_{i}^{-1} (u_{i}^{T} b) v_{i}, \quad \text{for } b \in \mathbf{B}$$
(12)

The pseudo-inverse (A^+) is also given by

$$A^{+} = \sum_{i=1}^{rank(A)} v_{i} \sigma_{i}^{-1} u_{i}^{T},$$
(13)

and the least squares solution \hat{x}_{ls} to the least squares problem is given

$$\hat{x}_{ls} = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i \tag{14}$$

The equation (14) shows that, it is the small singular values that magnify round-off errors, thus resulting in large errors in the solution. The sensitivity or stability of the solution x and x_{ls} to perturbations of A and B can be measured by the 2-norm, with condition number of A defined as:

$$cond(\mathbf{A}) = \|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2 = \frac{\sigma_l}{\sigma_s}$$
(15)

Thus, the condition number of A can be defined as the ratio of the largest singular values (σ_i) and the smallest singular values (σ_s) of A.

If the formulated problem above is regularized using truncated singular value decomposition, we obtain the result © 2013, IJMA. All Rights Reserved 122

Joseph Acquah & Ernest Danso-Addo / REGULARIZATION OF ILL-CONDITIONED LINEAR SYSTEMS/ IJMA- 4(7), July-2013. shown below.

i	σ_{i}	α	Rank(r)	max.error
1	$1.7954x10^{0}$		1	
2	$3.8028 x 10^{-01}$		2	
3	$4.4739 x 10^{-02}$		3	
4	3.7223×10^{-03}		4	
5	2.3309×10^{-04}		5	
6	$1.1163x10^{-05}$	$1.1163x10^{-05}$	6	$1.09x10^{-03}$
7	$4.0824 x 10^{-07}$	$4.0824x10^{-07}$	7	$1.63x10^{-04}$
8	$1.1229 x 10^{-08}$	$1.1229 x 10^{-08}$	8	$2.91x10^{-05}$
9	$2.2520x10^{-10}$	$2.2520x10^{-10}$	9	$3.57 x 10^{-06}$
10	$3.1113x10^{-12}$	$3.1113x10^{-12}$	10	$4.51x10^{-05}$
11	$2.6492 x 10^{-14}$	$2.6492 x 10^{-14}$	11	$1.19x10^{-03}$
12	$1.0772 x 10^{-16}$	$1.0772x10^{-16}$	12	$3.31x10^{-01}$

Table 2: Truncated Singular Values

7. PRECONDITIONING

The use of iterative methods for solving symmetric positive definite systems of linear equations, require some form of preconditioning M to improve the convergence of the solution by manipulating the spectrum of the coefficient matrix. Given the linear system Ax = B, we transform it into an equivalent system of the form MAx = MB, such that the conditioned number of MA is far less than that of A, that is $k(MA) \ll k(A)$. Basically, there are two types of preconditioner's, the left and the right preconditioner's, but for this study we will restrict our self to only the left preconditioner.

The Jacobi and Gauss-Seidel Preconditioner

The method of Jacobi and Gauss-Seidel for solving Ax = B, split the matrix A into A = L + D + U, where L is a lower triangular matrix, U is an upper triangular matrix and D a diagonal matrix. The Jacobi scheme is given by

$$x^{k+1} = -D^{-1}(U+L)x^{k} + D^{-1}B, \ k = 0,1,2,\dots$$
(16)

And the Gauss-Seidel scheme is also given by

$$x^{k+1} = -(L+D)^{-1}Ux^{k} + (L+D)^{-1}B$$
(17)

Both schemes converge to the solution if the matrix is *strictly diagonally dominant*. In the Jacobi iteration, the matrix D^{-1} is used to rescale all the non-diagonal entries of the matrix A, to obtain a good preconditioner, known as the **Jacobi Preconditioner**.

Thus, if A is ill-conditioned then $M = D^{-1}A$ is better conditioned than A. A Gauss-Seidel preconditioner can be used to solve the same problem. Here, the pre-conditioning matrix is lower triangular, and is defined as $M = (L + D)^{-1}$, from the Gauss-Seidel iteration. Solving the linear system with a Gauss-Seidel preconditioner is computationally expensive, but at times yields a better result than the Jacobi Preconditioner. Both the Jacobi and the Gauss-Seidel schemes are used for linear systems where the coefficient matrix is sparse, consist mainly of zeros. Such matrices occur in the numerical solution of boundary-value problems. The following examples illustrate the use of these preconditioner's.

8. A SIMPLE EXAMPLE

Consider the boundary-value problem

$$u_{xx} + u_{yy} = 0$$

© 2013, IJMA. All Rights Reserved

(18)

In the rectangle $R = \{(x, y) : 0 \le x \le 4, 0 \le y \le 4\},\$

where
$$u(x, y)$$
 denotes the temperature at the point $u(x, y)$, with boundary values $u(x,0) = 180$, $u(x,4) = 20$, for $0 < x < 4$
 $u(0, y) = 80$, $u(4, y) = 0$, for $0 < y < 4$

Central-difference approximation for u_{xx} and u_{yy} are given by

$$u_{xx} = \frac{u(x+h, y) - 2u(x, y) + u(x-h, y)}{h^2}$$
(19)

$$u_{yy} = \frac{u(x, y+h) - 2u(x, y) + u(x, y-h)}{h^2}$$
(20)

Then,

$$u_{xx}(x, y) + u_{yy}(x, y) \approx u(x + h, y) + u(x - h, y) + u(x, y + g) + u(x, y - g) - 4u(x, y) = 0$$
(21)

Choosing h = g = 1, corresponds to approximating the temperature u(x, y) at nine interior points in the rectangle shown in Figure 1.



Figure 1: :The Grid for the Example

Using the central difference formula for the second derivative, to approximate the partial derivatives at each of the nine interior points, results in the following linear systems. Thus u_{xx} and u_{yy} together with the *five-point formula* gives the following linear systems,

$$-4p_{1} + p_{2} + p_{4} = -100$$

$$p_{1} - 4p_{2} + p_{3} + p_{5} = -20$$

$$p_{2} - 4p_{3} + p_{6} = -20$$

$$p_{1} - 4p_{4} + p_{5} + p_{7} = -80$$

$$p_{2} + p_{4} - 4p_{5} + p_{6} + p_{8} = 0$$

$$p_{3} + p_{5} - 4p_{6} + p_{9} = 0$$

$$p_{4} - 4p_{7} + p_{8} = -260$$

$$p_{5} + p_{7} - 4p_{8} + p_{9} = -180$$

$$p_{6} + p_{8} - 4p_{9} = -180$$

(22)

where $p_1, p_2, ..., p_9$ are approximations of the temperature in the interior of the rectangle.

In matrix form, we have Ap = B, where A is a 9 x 9 symmetric positive definite matrix and p the vector solution.

$$\begin{bmatrix} -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -4 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & -4 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -4 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 \end{bmatrix} \begin{bmatrix} -100 \\ -20 \\ -20 \\ -80 \\ 0 \\ -80 \\ 0 \\ -80 \\$$

If we solve the linear system Ap = B, for *p* and regularized the solution using the Jacobi and Gauss-Seidel schemes, we obtain the solutions shown in figure 3.

X_J	X_{G}	X_{Jpred}	X_{GSpred}	Exact
55.6630	55.7141	55.7143	55.7143	55.7143
43.1459	43.2141	43.2143	43.2143	43.2143
27.0915	27.1427	27.1429	27.1429	27.1429
79.5744	79.6427	79.6429	79.6429	79.6429
69.8974	69.9998	70.0000	70.0000	70.0000
45.2887	45.3570	45.3571	45.3571	45.3571
112.8058	112.857	112.857	112.857	112.857
111.7173	111.785	111.786	111.786	111.786
84.2344	84.2856	84.2857	84.2857	84.2857

Table 3: Optimal Solution For Jacobi and Gauss-Seidel Preconditioner's

From Table, we conclude that an optimal solution is possible, when Boundary value problems are discretized into a linear system and solve using Preconditioning.

9. TIKHONOV REGULARIZATION METHOD

Regularization methods for least square problems are the most commonly used method for obtaining stable and smooth solution to rank deficient and ill-posed problems, [8]. In solving such problems, it is necessary to incorporate additional information as smoothness, continuity and the size of the residual to obtain the desired solution for x. Such additional information is then used as a side constraint to control the smoothness of the solution. The side constraint is usually of the form

$$C_{\lambda}(x) = \lambda \left\| Lx \right\|_{2}^{2}, \tag{24}$$

Where *L* is the identity matrix (I_n) or an $(n \ x \ p) \ x \ n$ discrete approximation of the p^{th} derivative operator. The side constraint gives a fair balance between minimizing $C_{\lambda}(x)$ and minimizing the residual norm $\|Ax - B\|_2^2$ instead of

giving us the solution Ax = B. The basic idea is that a regularized solution x should give a small residual and also be small in 2- norm to give a desired solution. One of the most important forms of regularization of ill-posed least squares problems is the **Tikhonov Regularization**. This method is often used to regularize ill-posed problems. It involves obtaining the exact or least squares solution of linear systems by minimizing the function

$$\phi_{\lambda}(x) = \left\| \mathbf{A}x = \mathbf{B} \right\|_{2}^{2},\tag{25}$$

subject to the side constraint ||Lx||, where the matrix L is either an identity matrix, a diagonal weighting matrix or an (n - p) x n discrete approximation of the p^{th} derivative operator. Since standard Algorithms normally fails to provide suitable solution to stabilize the system, the regularized solution x_{λ} , is defined as the minimizer of the weighted combination of the residual and the side constraint according to Wang and Linz (2003).

The minimized expression is

$$\phi_{\lambda}(x) = \|\mathbf{A}x - \mathbf{B}\|_{2}^{2} + \lambda \|Lx\|_{2}^{2},$$
(26)

where λ is greater than zero and is called a regularization parameter.

Regularization of Order Zero

If we let $L = I_n$, the minimizing function becomes

$$\phi_{\lambda}(x) = \|Ax - B\|_{2}^{2} + \lambda \|x\|_{2}^{2}$$
(27)

This signifies that it can be expressed as a balance between the quantities, $\|Ax - B\|_2^2$ and $\|x\|$. Here, the regularization parameter controls the weights given to minimization of the side constraint relative to the minimization of the residual norm. But we can show that the minimizing solution (x_{λ}) is given by the non-singular linear system as $(A'A + \lambda I)x_{\lambda} = A'B$. From

$$\phi_{\lambda}(x) = \|\mathbf{A}x - \mathbf{B}\|_{2}^{2} + \lambda \|x\|_{2}^{2},$$
(28)

It follows that

$$\phi_{\lambda} = (\mathbf{A}x)'(\mathbf{A}x) - (\mathbf{A}x)'\mathbf{B} - \mathbf{B}'\mathbf{B} + \lambda x'x$$
⁽²⁹⁾

But
$$(Ax)$$
'B and B' (Ax) are equal since they are scalars. So
 $\phi_{\lambda} = (Ax)'(Ax) - 2(Ax)'B + \lambda x'x + B'B,$
(30)

$$\phi_{\lambda} = x'(A'A)'x - 2A'x'B + \lambda x'x + B'B, \tag{31}$$

Differentiating the function $\phi_{\lambda}(x)$ for the minimizing solution x_{λ} , we obtain

$$\frac{\delta\phi_{\lambda}(x)}{\delta x}|x=x_{\lambda}=0$$
(32)

It implies that

$$\frac{\delta\phi_{\lambda}(x)}{\delta x}\Big|x = x_{\lambda} = 2\mathbf{A}'\mathbf{A}x_{\lambda} - 2\mathbf{A}'\mathbf{B} + 2\lambda x_{\lambda} = 0,$$
(33)

Or
$$(A'A + \lambda I)x_{\lambda} = A'B.$$
 (34)

Thus,

$$x_{\lambda} = (\mathbf{A}'\mathbf{A} + \lambda \mathbf{I})^{-1} (\mathbf{A}'\mathbf{B})$$
(35)

The regularization method above penalizes large components in the solution and is called regularization of Order Zero.

Regularization of Order One and Two

Regularization of Equation 27 in most cases dampens components that are large in magnitudes, since the component in a solution oscillates with moderate amplitudes. Such component are undesirable and may need a penalty term that is large for rapid change in the solution. This penalty term result in another form of regularization called "Order One". For this reason the penalty term is added to Equation 27 to obtain

$$\phi_{\lambda}(x) = \left\| \mathbf{A}x - \mathbf{B} \right\|_{2}^{2} + \lambda \sum_{i=2}^{n} \left(\left| x \right|_{i} - \left| x \right|_{i-1} \right)^{2}$$
(36)

The above expression is minimized by the solution of $(A^T A + \lambda L_1^T L_1)x_{\lambda} = A^T B$, where L_1 is an $(n - 1) \times n$ first derivative operator defined as

$$L_{1} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & 0 & 1 & -1 & \vdots \\ \vdots & \ddots & 1 & \ddots & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

For regularization of Order Two, the penalty is much stronger than that of Order One, and is based on minimizing the function

$$\phi_{\lambda}(x) = \left\| \mathbf{A}x - \mathbf{B} \right\|_{2}^{2} + \lambda \sum_{i=2}^{n} \left(\left| x \right|_{i+1} - 2 \left| x \right| + \left| x \right|_{i-1} \right)^{2}, \tag{37}$$

Which lead to the system $(A^T A + \lambda L_1^T L_1)x_{\lambda} = A^T B$. Here L_2 is an $(n-2) \times n$ second derivative operator defined as

$$L_2 = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{bmatrix}$$

1

The second derivative operator L_2 also helps us to compute x_{λ} for Order Two regularization using MATLAB or OCTAVE Script files for regularization.

10. APPLICATION TO THE SOLUTION OF FREDHOLM INTEGRAL EQUATION OF THE FIRST KIND

Consider Fredholm integral equation of the first kind given by

$$g(s) = \int_{0}^{\infty} K(s,t)x(t)dt \quad \text{for} \quad s \in [0,1],$$
(38)

where the right-hand side vector g(s) and the kernel K are unknown functions, while x(t) is an unknown solution [4].

If the integral equation is discretized using the Quadrature (or Galerkin) methods, a quadrature rule with abscissas $t_1, t_2, t_3, \dots, t_n$ and corresponding weights $w_1, w_2, w_3, \dots, w_n$ is obtained as an approximate to the integral equation. The integral equation then becomes:

$$\int_{0}^{1} K(s,t)x(t)dt = \sum_{j=1}^{n} w_{j}K(s,t)x(t_{j})$$
(39)

If the approximation is applied to an integral equation with m distinct point values $s_1, s_2, s_3, \dots, s_m$, we obtain an $m \ge n$ matrix A given by $a_{ij} = w_j k(s_i, t_j)$ and a right hand side vector B, also by $B_i = g(s_i)$. The equation Ax = B becomes:

$$g(s_i) = \sum_{j=1}^n w_j k(s_i, t_j) x(t_j)$$
(40)

11. A SAMPLE PROBLEM

Consider the integral equation

$$\int_{0}^{1} e^{(s+1)t} x(t) dt = \frac{e^{s+1} - 1}{s+1}$$
(41)

To solve the equation numerically, we discretized the equation into a linear system $A_{ij}x_j = B_i$ or $\sum w_j k(s_i, t_j)x_j = g(s_i)$, by using equation (40), with an n-point composite trapezoidal rule and with uniformly spaced quadrature points. From (40), the kernel of the integral equation in (40) is $k(t_i, t_j) = e^{(t_i+1)t_j}$, with its weight function as w_j , the computed solution as x_j , and the function $g(t_i)$ also given by

$$g(t_i) = \frac{e^{(t_i+1)}-1}{t_i+1},$$

where $s = t_i$ and $t = t_j$ for i = 1, 2, 3, ..., n and j = 1, 2, 3, ..., n

Thus, for a five point composite trapezoidal rule, the integral limits from 0 to 1 is divided into a number of strips of size n = 4, to obtain the five points, using equation (40). While for small values of n, the computed solution show some resemblance to the exact solution, the error increases rapidly with n. Thus no value of x approximates to the exact solution with acceptable accuracy. The Table 4 shows the computed and the exact solutions for selected values of n.

Point s(t)	<i>n</i> = 4	<i>n</i> = 8	<i>n</i> = 16	Exact
0.0000	0.5884	0.5306	$3.1775e^4$	1.0000
0.1250	_	1.7584	$-5.9354e^{3}$	1.0000
0.2500	1.4734	-0.5601	$1.3437e^{3}$	1.0000
0.3750	_	3.4871	$3.9364e^{3}$	1.0000
0.5000	0.4737	-1.8604	$3.0043e^{3}$	1.0000
0.6250	_	3.3715	$6.0280e^{3}$	1.0000
0.7500	1.4539	-0.4432	$1.0349e^4$	1.0000
0.8750	_	1.7075	$7.1782e^{3}$	1.0000
1.0000	0.6093	0.5480	$-1.9681e^{3}$	1.0000

Table 4: Computed Solution for Selected Value of n.

The reason for this anomaly is quite clear when we look at the condition numbers of some selected values of n, shown in the Table 5 below.

Table 5: Condition Numbers for Selected Values of n.

n	Condition Number	
4	$1.1649 x 10^5$	
8	$4.5787 x 10^7$	
16	5.7636×10^{15}	

From the table, as n increases from 4 onward, we expect the discretization error to reduce, that is the accuracy with which (41) represents (40) gets better. Unfortunately this is not so, the condition numbers rather increases rapidly destroying any gain from the more accurate discretization. This makes it quite clear that any computations with n greater than sixteen will make the solution worse.

12. APPLYING TIKHONOV REGULARIZATION IN SOLVING THE SAMPLE PROBLEM

Generally, when we solve ill-posed problems by discretizing the equation, the matrix we get is inherently illconditioned, and the condition numbers increases rapidly with n. Thus, there is the need to apply some form of regularization in order to obtain reasonable approximate solution.

Joseph Acquah & Ernest Danso-Addo / REGULARIZATION OF ILL-CONDITIONED LINEAR SYSTEMS/ IJMA- 4(7), July-2013.

When Order Zero, One and Two Tikhonov regularization was applied to our sample problem for the special case n = 16, the solutions for Order One and Two converges approximately to the exact solution, but that of Order Zero shows little convergence to the exact solution. A detailed solution for Order Zero, One and Two, for a given range of values of the regularization parameter λ was computed and the best for each order selected. These convergent solutions are not the same as the optimal solution; rather it gives us a fare idea of the optimal solution. The respective convergent solutions are shown in Table 6.

Order Zero	Order One	Order Two	Exact
0.40206	0.99886	0.9984229	1.0000
0.84232	0.99886	0.9984659	1.0000
0.88038	0.99886	0.9985090	1.0000
0.91211	0.99886	0.9985521	1.0000
0.95428	0.99886	0.9985951	1.0000
0.98904	0.99886	0.9986382	1.0000
1.00440	0.99886	0.9986812	1.0000
1.05060	0.99886	0.9987243	1.0000
1.07560	0.99886	0.9987673	1.0000
1.09520	0.99886	0.9988104	1.0000
1.10810	0.99886	0.9988535	1.0000
1.11260	0.99886	0.9988965	1.0000
1.10710	0.99886	0.9989396	1.0000
1.08930	0.99886	0.9989826	1.0000
1.05670	0.99886	0.9990257	1.0000
1.01090	0.99886	0.9990687	1.0000
0.46762	0.99886	0.9991118	1.0000

Table 6: Convergent Regularized Solution for n = 16

13. DETERMINATION OF OPTIMAL REGULARIZATION PARAMETER AND SOLUTION

In determining the optimal regularization parameter λ corresponding to the optimal solution, two methods can be applied. The first is by inspection and the second by the use of the L-Curve approach. For the method by inspection, we compute the residual norm, the solution norm and inspect the result for a range of values of the parameter λ . The behavior of the residual norm gives a good indication of a proper choice of λ . If the residual norm increases or decreases steadily with λ , until a critical value λ_0 , where further increases or decreases have no effect on the

residual norm, then λ_0 is chosen as the optimal regularization parameter and its corresponding value becomes the optimal solution. But the method by inspection, is a times very difficult to use in estimating the optimal solution.

However, it is more appropriate to use the L-Curve method coupled with the method of inspection to determine the critical value for the optimal solution. The two methods together help us to identify clearly all the parameter points. Here, we plot the solution norm against the residual norm to obtain the L-Curve. The optimal regularization parameter λ_0 is usually identified at the sharp corner of the L-Curve. Normally, we make use of the points concentrated at the sharp corner of the L-Curve to determine the optimal regularization parameter λ_0 . The optimal solution corresponding to the optimal regularization parameter λ_0 is obtained from the column regularized solutions computed. The first column corresponds to $\lambda_1 = 10^{-16}$, the second $\lambda_2 = 10^{-15}$, and the rest continues in that order. The L-Curve for Order Zero, One and Two for the sample problem are shown in the figures 2, 3, and 4 respectively.



Figure 2: L-Curve for Order Zero Regularization

For Order Zero regularization, the regularization parameters concentrated at the sharp corner of the L-Curve by inspection, are $\lambda_4 = 10^{-13}$ and $\lambda_5 = 10^{-12}$. The regularized solutions corresponding to the parameters $\lambda_4 = 10^{-13}$ and $\lambda_5 = 10^{-12}$, do approximate a little to the exact solution, with λ_5 better conditioned than λ_4 . Hence, order zero regularization for this problem has an optimal solution with regularization parameter $\lambda_5 = 10^{-12}$.



Regularization for Order One

Figure 3: L-Curve for Order One Regularization

For Order One regularization, the regularization parameters concentrated at the sharp corner of the L-Curve by inspection, are $\lambda_7 = 10^{-10}$ and $\lambda_8 = 10^{-9}$. The regularized solutions corresponding to the parameters $\lambda_7 = 10^{-10}$ and $\lambda_8 = 10^{-9}$, do approximate to the exact solutions better compared to that of Order Zero. Therefore, Order One regularization in this case, has an optimal solution with a regularization parameter of $\lambda_8 = 10^{-9}$.



Figure 4: L-Curve for Order Two Regularization

For Order Two regularization, the regularization parameters concentrated at the corner of the L-Curve are $\lambda_{12} = 10^{-5}$, $\lambda_{13} = 10^{-4}$ and $\lambda_{14} = 10^{-3}$ respectively. All the regularized solutions corresponding to the parameters $\lambda_{12} = 10^{-5}$, $\lambda_{13} = 10^{-4}$ and $\lambda_{14} = 10^{-3}$ approximate to the exact solution, but the entries of the solution for $\lambda_{13} = 10^{-4}$ are well skewed and better approximate to the exact solution compared to the solutions of $\lambda_{12} = 10^{-5}$ and $\lambda_{14} = 10^{-3}$. Therefore, Order Two regularization in this case, has an optimal solution with regularization parameter $\lambda_{13} = 10^{-4}$.

14. THE OPTIMAL SOLUTION

Comparing the optimal solutions for Order Zero, One and Two regularization with the exact solution, it is observe that, the optimal solution for Order Two are well skewed and approximate better to the exact solution. The optimal regularization parameter for Order Two is $\lambda_{13} = 10^{-4}$ and the corresponding optimal solution ($x_{\lambda_{13}}$) is shown in Table 7.

Exact Solution (x)	Regularized Solution \hat{x} at $\lambda_{13} = 10^{-4}$	Unregularized Solution	
1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000	$x \ at \ \lambda_{13} = 10$ 0.9984226 0.9984656 0.9985087 0.9985947 0.9986377 0.9986807 0.9987237 0.9987667 0.9988097 0.9988957 0.9988957 0.9988957 0.9989387 0.9989316 0.9990245 0.9900675	$\begin{array}{r} 3.1775e^4 \\ -2.0710e^4 \\ -5.9354e^3 \\ -8.0580e^{-1} \\ 1.3437e^3 \\ 1.9390e^4 \\ 3.9364e^3 \\ -1.2034e^4 \\ 3.0043e^3 \\ -1.3519e^4 \\ 6.0280e^3 \\ -1.4620e^1 \\ 1.0349e^4 \\ -1.3898e^4 \\ 7.1782e^3 \end{array}$	
1.0000000	0.9991094	$-2.5030e^{-1}$ -1.9681 e^{3}	

Table 7: Optimal Solution

Joseph Acquah & Ernest Danso-Addo / REGULARIZATION OF ILL-CONDITIONED LINEAR SYSTEMS/ IJMA- 4(7), July-2013.

In conclusion, the optimal solution for the problem approximate accurately to the exact solution, compared to the unregularized solution which shows no resemblance to the exact solution. The maximum error in the computed solution reduced drastically to $1.5774e^{-1}$ in the optimal solution.

15. CONCLUSION

In this paper, we studied various methods for solving ill-conditioned linear systems, using the Hilbert system as a prototype. The regularized solutions of Order One and Two for our test problem gave an accuracy of about 3 digits of precision, with parameter values $\lambda = 10^{-9}$ and $\lambda = 10^{-4}$ respectively. Order Zero shows no accuracy in digits precision of the regularized solutions. The L-Curve was used to determine the optimal regularization parameter. The optimal regularization parameter corresponding to the optimal solution was determined at the corner part of the L-Curve. The L-Curve for Order Two regularization gave us the optimal solution with a regularization parameter value $\lambda_{13} = 10^{-4}$. In conclusion, we justify that in the absence of computational errors and with proper regularization techniques, a convergent discretization error leads to a convergent solution.

REFERENCES

- 1. P. C. Hansen, Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of *Linear Inversion*. *Philadelphia: SIAM, 1998*.
- 2. P. Linz and R. L. Wang, Exploring Numerical Methods: An Introduction to Scientific Computing Using MATLAB. New York: Jones and Bartlett Publishers, 2003
- 3. Groetsch N. (1984): The Theory of Tikhonov Regularization for Fredholm Integral Equations of the First Kind.London: Pitman.
- 4. Wing, G. M. and Zahrt. J. D. (1991). A prime on integral equations of the first kind. Siam Review, Philadelphia, pp. 191-204
- 5. Benyah , F. and Jennings, L. (1998). Regularization in Optimal Control Computations. Anchorage, Alaska, International Symposium on Intelligent Automation and Control, World Automation Congress, pp. 78-135.
- 6. Maziar Salahi, (2009). On Il l-Conditioned Linear systems and Regularized GMRES *Method*. Journal of Applied Mathematics and Information Science, vol.3(1), pp.45-51.
- 7. Arnold Neumaier, (1998). Solving Ill-Conditioned and Singular Linear systems: A Tutorial on Regularization Siam Review, vol.40, No.3, pp. 636-666.
- 8. Hansen, P. C. (1987). Truncated Singular Value as a Method for Regularization. BIT Review 27, pp. 534-553.
- 9. Benyah, F. and Jenning, L. (1998) .The L-Curve in the Regularization of Optimal Control *Computations, J. Austral. Math. So c., Series B E40, pp. 138-172.*

Source of support: Nil, Conflict of interest: None Declared