Available online at www.sciencedirect.com

ScienceDirect

Applied Mathematical Modelling 00 (2014) 1–16

Applied Mathematical Modelling

An efficient method to set up a Lanczos based preconditioner for discrete ill-posed problems

Shervan Erfani[a], Ali Tavakoli[a,*], Davod Khojasteh Salkuyeh[b]

[a]*Mathematics Department, Vali-e-Asr University of Rafsanjan, Iran.*
[b]*Faculty of Mathematical Sciences, University of Guilan, P.O. Box 1914, Rasht, Iran*

Abstract

Rezghi and Hosseini [M. Rezghi and S. M. Hosseini, Computing 88(2010), 79-96] presented a Lanczos based preconditioner for discrete ill-posed problems. Their preconditioner is constructed by using few steps (e.g., k) of the Lanczos bidiagonalization and corresponding computed singular values and right Lanczos vectors. In this article, we propose an efficient method to set up such preconditioner. Some numerical examples are given to show the effectiveness of the method.

Keywords: Lanczos bidiagonalization, Ill-posed problems, Preconitioner, Regularization, Singular value decomposition.
2000 MSC: 35J65, 58J90

1. Introduction

We consider large scale discrete ill-posed problems, i.e., linear systems of equations of the from

$$Ax = b, \qquad A \in \mathbb{R}^{n\times n}, \quad x, b \in \mathbb{R}^n. \tag{1}$$

The coefficient matrix A is of full-rank and typically ill-conditioned, the right-hand side vector b is perturbed by an error such that $b = \hat{b} + e$, which may occur by measurement or discretization errors. Here e and $\hat{b}$ denote the Gaussian noise and unknown error-free right-hand side vectors, respectively. These problems typically arise, for instance, in discretizing linear ill-posed problems, such as Fredholm integral equation of the first kind with a smooth kernel. Also in these problems the two following criteria are satisfied:

(1) The singular values of A decay gradually to zero with no particular gap in the spectrum.

(2) The ratio between the largest and the smallest nonzero singular values is large [7].

*Corresponding author
Email addresses: sh.erfani@stu.vru.ac.ir (Shervan Erfani), tavakoli@mail.vru.ac.ir (Ali Tavakoli), khojasteh@guialn.ac.ir, salkuyeh@gmail.com (Davod Khojasteh Salkuyeh)
URL: tavakoli.vru.ac.ir (Ali Tavakoli)

We would like to obtain an approximate solution $\hat{x}$ of the linear system $A\hat{x} = \hat{b}$ with error-free unknown right-hand side. Since the matrix A is severely ill-conditioned, the solution of system (1) typically is not a significant approximation of $\hat{x}$ even if the noise e is small.

So, we use the regularization methods to determine a solution that approximates the exact solution $\hat{x}$. The conjugate gradient method applied to normal equations (CGLS) is a well-known iterative method for solving ill-posed problems. Since, iterative methods, such as CGLS, have low rate of convergence, it is possible to speed up the convergence by a suitable preconditioner [15].

All regularization methods use one or more regularization parameters specific to the regularization method that controls the amount of the stabilization imposed on the solution, and in most cases it is necessary to choose this parameter from given the problem and the given set of data. A large regularization parameter makes a new well-conditioned problem, but its solution may be far from the exact solution. A small regularization parameter generally yields a solution very close to the noise-contaminated exact solution of (1), and hence its distance from the noise-free solution also can be large. Thus, one can choose a regularization parameter to balance the error due to noise with the error due to regularization. A good choice of regularization parameter is clearly crucial to determine useful approximate solution for ill-posed problems [12].

In [15], Rezghi and Hosseini proposed a new preconditioner for discrete ill-posed problems which is produced by a few steps of Lanczos bidiagonalization process. The computation of their preconditioner depends on the large singular values of A. In this paper, we propose a new and efficient numerical method to estimate the large singular values of the matrix A and construct the new form of Lanczos based preconditioner, which can be obtained in a very small number of operations.

The paper is organized as follows. In Section 2, we study Fredholm integral equation of the first kind and introduce the quadrature rule and Galerkin method for discretization of these equations. Also, we study the treatment of the singular values on the matrix resulted by discretization. In Section 3, we define Lanczos bidiagonalization for constructing Least squares via Lanczos bidiagonalization (LSQR) method and preconditioner to approximate the solution of the ill-posed problems. In Section 4, we introduce a new method for constructing the new form of Lanczos based preconditioner and finding regularization parameter based on large singular values of the singular spectrum. Finally, in Section 5 some numerical experiments are given.

2. Fredholm Integral Equation of the First Kind

2.1. The Singular Value Expansion

A classical example in linear ill-posed problems is a Fredholm integral equation of the first kind with a square integrable kernel

$$\int_a^b K(s,t)f(t)dt = g(s), \qquad c \le s \le d, \tag{2}$$

where the right-hand side g and kernel K are known functions and f is an unknown solution. Fredholm equation of the first kind can be classified into two categories. The first one arises if the kernel function $K(s,t)$ is smooth. In this case, Fredholm equation is very often extremely ill-conditioned, i.e., small changes in the data cause huge changes in the unknown function f. That means the solution f is extremely sensitive to small changes in $g(s)$ and so special numerical techniques called regularization methods are required. In the second category, the kernel $k(s,t)$ is a singular function [8].

The superior analytical tool for analysis of the first-kind Fredholm integral equation (2) with square integrable kernels is the singular value expansion (SVE) of the kernel. A kernel K is called square integrable if the norm

$$\|K\|^2 = \int_a^b \int_c^d K(s,t)^2 dsdt,$$

is bounded. By means of the SVE, any square integrable kernel K can be written as the following infinite sum [7]:

$$K(s,t) = \sum_{i=1}^{\infty} \mu_i u_i(s) v_i(t).$$

The functions u_i and v_i are said to be the singular functions of K. They are orthonormal with respect to the usual inner product, i.e.,

$$(u_i, u_j) = (v_i, v_j) = \begin{cases} 1, & \text{if} \quad i = j, \\ 0, & \text{if} \quad i \neq j, \end{cases}$$

where $(.,.)$ is defined by

$$(\phi, \psi) = \int_a^b \phi(t)\psi(t)dt.$$

The numbers μ_i's are the singular values of K, which are nonnegative and can always be ordered in non-increasing order so that

$$\mu_1 \geq \mu_2 \geq \mu_3 \geq \cdots \geq 0.$$

And, if there exists a positive real number α such that the singular values satisfy $\mu_i = \mathcal{O}(i^{-\alpha})$, then α is called the degree of ill-posedness, and problems is characterized as mildly or moderately ill-posed if $\alpha \leq 1$ or $\alpha > 1$, respectively. On the other hand, if $\mu_i = \mathcal{O}(e^{-\alpha i})$, then the problem is termed severely ill-posed (see [7]). By the definition of $\|K\|$, we have

$$\begin{aligned} \|K\|^2 &= \int_a^b \int_c^d \sum_{i=1}^{\infty} \mu_i u_i(s) v_i(t) \sum_{j=1}^{\infty} \mu_j u_j(s) v_j(t) ds dt \\ &= \int_a^b \int_c^d \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mu_i \mu_j u_i(s) u_j(s) v_i(t) v_j(t) ds dt. \end{aligned} \tag{3}$$

So, we have

$$\| K \|^2 = \sum_{i=1}^{\infty} \mu_i^2 < \infty.$$

Since the sequence $\{\mu_i\}_{i=1}^{\infty}$ is nonnegative and $\sum_{i=1}^{r} \mu_i^2$ for all r is bounded, then $\sum_{i=1}^{\infty} \mu_i^2$ is convergent. Also, we have $\sum_{i=1}^{\infty} i^{-\alpha}$ is convergent if $\alpha > 1$ and divergent if $\alpha \leq 1$.
This shows that the μ_i^2 must decay faster than i^{-1}. In general, if K is a continuous partial derivatives of order p then μ_i is approximately $\mathcal{O}(i^{-p-\frac{1}{2}})$ [11, 17]. So, we conclude that if the kernel K is smoother, then the singular values μ_i's decay to zero rapidly. This yields an increase in the number of small singular values.

2.2. Discretization

In order to solve Fredholm integral equation of the first kind we have to discretize it. There are essentially two main classes of methods, namely, quadrature and Galerkin methods. In the quadrature method, a quadrature rule like the midpoint rule, the trapezoidal rule and the Simpson rule, is used. The discretized form of (2) reads

$$\sum_{j=1}^{n} \omega_j K(s_i, t_j) f(t_j) = g(s_i), \qquad i = 1, 2, \ldots, n,$$

with the weights $\{\omega_j\}_{j=1}^{n}$ and the nodes $\{t_i\}_{i=1}^{n}$. So, Eq. (2) is approximated by the $n \times n$ linear system of equations

$$Ax = b,$$

where A is the $n \times n$ matrix with the entries $a_{ij} = \omega_j K(s_i, t_j)$, x is the n-vector

$$x = [f(t_1), f(t_2), \ldots, f(t_n)]^T,$$

and b is the n-vector

$$b = [g(s_1), g(s_2), \ldots, g(s_n)]^T.$$

In the Galerkin method, one can choose two different orthonormal basis functions, $\{\psi_i\}_{i=1,\ldots,n}$ and $\{\varphi_j\}_{j=1,\ldots,n}$, to expand the functions g and f respectively, by

$$g(s) = \sum_{i=1}^{n} g_i \psi_i(s) \quad \text{and} \quad f(t) = \sum_{j=1}^{n} f_j \varphi_j(t).$$

Thus,

$$g(s) = \sum_{j=1}^{n} f_j \int_a^b K(s,t)\varphi_j(t)dt,$$

so that,

$$g_i = \sum_{j=1}^{n} f_j \int_a^b \int_c^d K(s,t)\psi_i(s)\varphi_j(t)dsdt.$$

The system may be written in the matrix form as $Ax = b$, where

$$a_{ij} = \int_a^b \int_c^d K(s,t)\psi_i(s)\varphi_j(t)dsdt,$$

with $b = [g_1, g_2, \ldots, g_n]^T$ and $x = [f_1, f_2, \ldots, f_n]^T$.

2.3. The Singular Value Decomposition

Let $A \in \mathbb{R}^{n\times n}$. Then, the SVD of A is a decomposition of the from

$$A = U\Sigma V^T = \sum_{i=1}^{n} u_i \sigma_i v_i^T,$$

where $U = (u_1, u_2, \ldots, u_n)$ and $V = (v_1, v_2, \ldots, v_n)$ are unitary matrices, i.e., $U^TU = V^TV = I_n$, and where $\Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$ has non-negative diagonal entries appearing in non-increasing order such that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n.$$

The numbers $\sigma_i, i = 1, \ldots, n$ are the singular values of A. In continuation , we describe that an increase in the dimensions of A will only increase the number of small singular values. To this end, let $\mu_1, \mu_2, \ldots, \mu_k$ be the large singular values of the kernel K. As we said in the last subsection, if K is a smooth function, there exists a few large singular values μ_i that yield $k \ll n$. We have (see [10]):

$$\sigma_1 \approx \mu_1 \gg \sigma_2 \approx \mu_2 \gg \ldots \gg \sigma_k \approx \mu_k \gg 0.$$

On the other hand, since μ_i's are fixed, then the large singular values σ_i would be unchanged whenever the dimension of the matrix A is increased.
In connection with discrete ill-posed problems, two characteristic features of SVD are very often found

(1) The singular value σ_i decay gradually to zero with no particular gap in the spectrum. An increase of the dimensions of A will increase the number of small singular values.

(2) The left and right singular vectors u_i and v_i tend to have more sign changes in their elements as the index i increases (i.e., as σ_i's decreases).

Both features are consequences of the fact that the SVD of A is closely related to SVE of the underlying kernel K. In fact, in the singular values σ_i of A denote to many cases approximations of the singular values μ_i of K[2].
Let us consider, for example, the test problem **derive2(100,1)** in [8]. The kernel K of the integral equation (2) is Green's function for the second derivative:

$$K(s,t) = \begin{cases} s(t-1), & s \leq t, \\ t(s-1), & s \geq t, \end{cases}$$

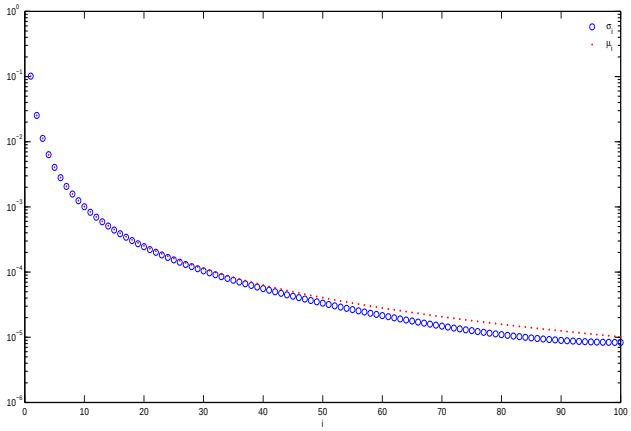

Figure 1. A comparison between σ_i and μ_i for *derive* test problem.

and both integration intervals are $[0, 1]$. The singular values and functions are given by (see for example [2])

$$\begin{cases} \mu_i = (i\pi)^{-2}, \\ u_i(s) = \pm\sqrt{2}\ \sin(i\pi s), \qquad i = 1, 2, \ldots, \\ v_i(t) = \mp\sqrt{2}\ \sin(i\pi t). \end{cases} \tag{4}$$

Since the singular values are proportional to i^{-2}, The problem is moderately ill-posed. Figure 1 shows that the singular values σ_i of discretized matrix of the test problem **deriv2(100,1)** with dimension 100. It is clear from this figure that the singular values of matrix A and the corresponding μ_i of kernel K are almost the same.

Since A is a full rank matrix, then its inverse is given by

$$A^{-1} = \sum_{i=1}^{n} v_i \sigma_i^{-1} u_i^T,$$

and therefore the solution of $Ax = b$ is

$$x = A^{-1}b = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i. \tag{5}$$

Moreover, since there are discretization errors or linear approximation errors, the discrete ill-posed problems always contain perturbations and error components created in tension all singular vectors of A. So, we cannot compute a stabilized solution. By (5), the errors in the vector b are amplified with coefficient σ_i^{-1}. If σ_i is close to zero then the solution is affected by the error e in the vector b,

$$b = \hat{b} + e, \quad \|\hat{b}\| \geq \|e\|.$$

In this case, we have

$$x = V\Sigma^{-1}U^T b = V\Sigma^{-1}U^T\hat{b} + V\Sigma^{-1}U^T e = \hat{x} + x_e,$$

where $\hat{x} = V\Sigma^{-1}U^T\hat{b}$ and $x_e = V\Sigma^{-1}U^T e$. If the singular values decay gradually to zero, then the term x_e overcomes the exact solution. Therefore, Fourier coefficients $|u_i^T b|$ corresponding to small singular values with lower rate tend to zero. Hence, the terms corresponding to small singular values overshadow the solution. We should use the regularization methods to subtract or eliminate the solution corresponding to small singular values.

3. Lanczos Bidiagonalization

For a given matrix $A \in \mathbb{R}^{n\times n}$, the Lanczos Bidiagonalization algorithm generates two orthogonal matrices U_n and V_n such that $U_n^T A V_n = B_n$, where B_n is a real lower bidiagonal matrix. This algorithm is described as follows:

Algorithm 1. `Lanczos Bidiagonalization`

1. `Let` $\beta_1 := \|b\|_2$, $u_1 := b/\beta_1$ `and` $v_0 := 0$
2. `For` $i = 1$ `to` n `do`
3. $p_i := A^T u_i - \beta_i v_{i-1}$
4. $\alpha_i := \|p_i\|_2$
5. $v_i := p_i/\alpha_i$
6. $q_i := Av_i - \alpha_i u_i$
7. $\beta_{i+1} := \|q_i\|_2$
8. $u_{i+1} := q_i/\beta_{i+1}$
9. `Endfor`

The vectors u_i and v_i produced by Lanczos Bidiagonalization algorithm are called the left and right Lanczos vectors, respectively. Denoting $U_k = [u_1, u_2, \ldots, u_k]$ and $V_k = [v_1, v_2, \ldots, v_k]$, the following relations can be established:

(1) $b = \beta_1 u_1 = \beta_1 U_{k+1} e_1$,
(2) $AV_k = U_{k+1} B_k$,
(3) $A^T U_{k+1} = V_k B_k^T + \alpha_{k+1} v_{k+1} e_{k+1}^T$,
(4) $V_k^T V_k = I_k$,
(5) $U_{k+1}^T U_{k+1} = I_{k+1}$,

where I_j denotes the identity matrix of order j, e_i is the ith unit vector and

$$B_k = \begin{pmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \beta_3 & \ddots & \\ & & \ddots & \alpha_k \\ & & & \beta_{k+1} \end{pmatrix}.$$

Now, suppose we want to solve

$$\min_{x\in S} \| b - Ax \|_2,$$

where S denotes the k-dimensional subspace spanned by the first k Lanczos vectors v_i, $i = 1, \ldots, n$. The solution which we seek is of the from $x^{(k)} = V_k y^{(k)}$ for some vectors $y^{(k)} \in \mathbb{R}^k$. Let $r^{(k)} = b - Ax^{(k)}$ be the corresponding residual. From the above relations, we get

$$r^{(k)} = \beta_1 u_1 - AV_k y^{(k)} = U_{k+1}(\beta_1 e_1 - B_k y^{(k)}).$$

Since the columns of U_{k+1} are orthonormal, we have

$$\|r^{(k)}\|_2 = \|\beta_1 e_1 - B_k y^{(k)}\|_2.$$

In the step k of the LSQR algorithm we are going to solve

$$\min_{y^{(k)}} \|\beta_1 e_1 - B_k y^{(k)}\|_2.$$

Now, we write the exact SVD of the computed B_k as

$$B_k = H_k \Gamma Q_k^T = \sum_{i=1}^{k} h_i \gamma_i q_i^T .$$

Then we obtain

$$x^{(k)} = \beta_1 V_k \sum_{i=1}^{k} \frac{h_{1i}}{\gamma_i} q_i, \tag{6}$$

where k is typically small.

The best value of k can be computed by one of the following four ways:

(1) In [15] the parameter k is considered as the smallest integer for which

$$\sigma_k < \tau \sigma_1, \tag{7}$$

where τ is the square root of the machine precision. Here, σ_i's are the singular values of B_k [15].

(2) **Generalized Cross-Validation (GCV)**
The parameter k is chosen by GCV method that minimizes the following function (see [12]):

$$\psi_k = \frac{\| \beta_1 e_1 - B_k y^{(k)} \|_2^2}{(n-k)^2}$$

(3) **L-Curve**
To determine the L-curve associated with LSQR, estimates of $\| r^{(k)} \|$ and $\| x^{(k)} \|$ are needed for several values of k. In the method, the corner of the L-curve gives a good balance of the solution size and residual size, so the parameter k can be chosen by the corner this curve (see [12]).

(4) **Discrete Picard Condition**
A standard tool for analyzing the discrete ill-posed problems is the discrete Picard plot, which is a plot of the quantities σ_i, $\mid u_i^T b \mid$ and $\frac{|u_i^T b|}{\sigma_i}$ that arise in (5). In order to derive a meaningful regularization solution in a discrete ill-posed problem, it must satisfies the discrete picard condition, i.e the Fourier coefficient $\mid u_i^T b \mid$ on the average should decay to zero faster than the singular values σ_i [8]. In this paper, we use the discrete Picard condition to find the parameter k.

Figure 2 shows plots of the first 200 singular values, Fourier coefficients $\mid u_i^T b \mid$, and coefficients $\frac{|u_i^T b|}{\sigma_i}$ of the solution for perturbed problem *shaw* [8] with Gaussian noise in which $\frac{\|e\|}{\|\hat{b}\|} = 1.10^{-3}$ and $\frac{\|e\|}{\|\hat{b}\|} = 1.10^{-1}$. As we observe, one can choose $k = 8$ and $k = 7$ for *shaw* problem with $\frac{\|e\|}{\|\hat{b}\|} = 1.10^{-3}$ and $\frac{\|e\|}{\|\hat{b}\|} = 1.10^{-1}$, respectively. We moreover see that the different perturbations contain a little change in the value of parameter k.

In [15], a new regularized preconditioner obtained by k steps of Lanczos bidiagonalization for discrete ill-posed problems has been introduced. The construction of this preconditioner is not based on any particular structure of the matrices. The matrix

$$M = V_k (B_k^T B_k)^{-1} V_k^T + (I - V_k V_k^T) \in \mathbb{R}^{n \times n}, \tag{8}$$

clusters approximately the large singular values around 1 and leaves the others unchanged.The idea here is based on the fact that the preconditioner is constructed by $k \ll n$ steps of Lanczos bidiagonalization.

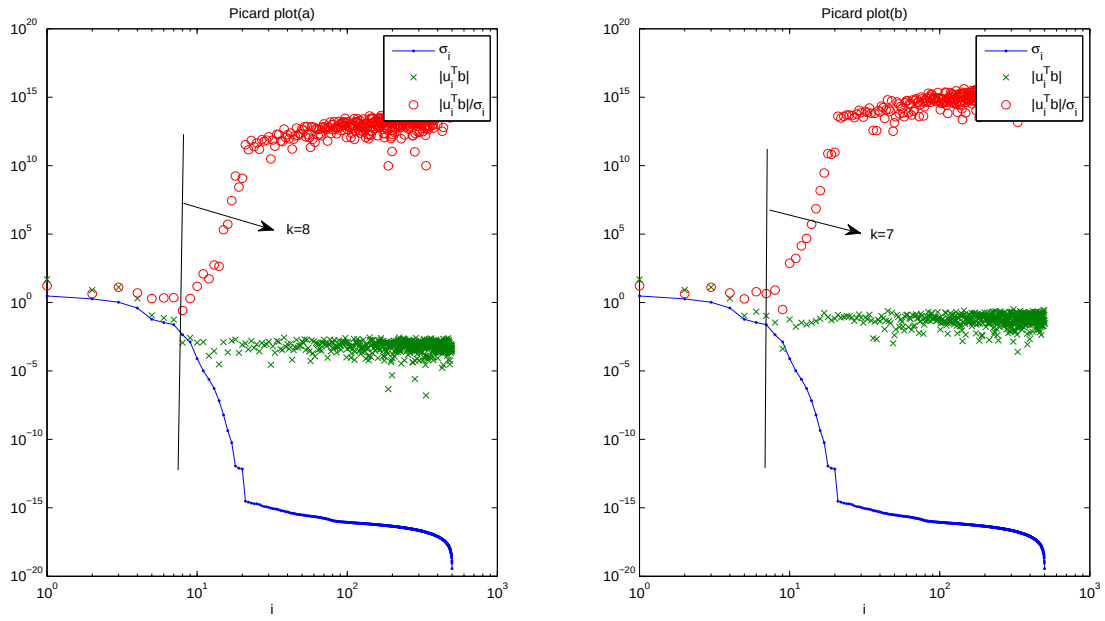

Figure 2. The first 200 singular values σ_i , Fourier coefficients $\mid u_i^T b \mid$, and coefficients $\frac{|u_i^T b|}{\sigma_i}$ for the *shaw* test problem with two perturbations $\frac{\|e\|}{\|\bar{b}\|} = 1.10^{-3}$ (a) and $\frac{\|e\|}{\|\bar{b}\|} = 1.10^{-1}$ (b).

4. An effective method for computing optimal Lanczos based preconditioner

We know, the matrix A can be produced by discretization of Fredholm integral equation of the first kind. Solving these equations by methods like Galerkin or quadrature rules usually produce an ill-posed system $Ax = b$ that yields a huge condition number of A. The solution of the system $Ax = b$ converges well, if the dimensions of A are sufficiently large (see for instance [3]). However, in this case, the singular values of A generated by SVD are not exact, and hence make the selection of parameter k difficult.

In the sequel, we introduce a criterion for choosing a suitable regularization parameter k. To do so, we discretize the Fredholm integral equation on a coarse grid that produces the system $\bar{A}\bar{x} = \bar{b}$. Then the matrix $\bar{A}$ resulted from this kind of discretization has less dimensions than those of A (e.g, half of A) and its singular values can be absolutely computed more easily than the original matrix. On the other hand, as we mentioned in Section 2, an increase in the dimensions of A will increase the number of small singular values. Thus, the number of *large singular values* (i.e., k) for A and $\bar{A}$ are almost the same. Moreover the number k (for the matrix $\bar{A}$) can be derived via the methods listed in Section 3.

All we need is to compute the right Lanczos vectors V_k. To do this, we first note that such vectors are available for $\bar{A}$. Let $\bar{v} = (\bar{v}^1, \bar{v}^2, \ldots, \bar{v}^m)^T$ be a right Lanczos vector corresponding to a coarse grid. We use an interpolation (prolongation) operator I_k as follows:

$$I_k \bar{v} = w,$$

where

$$w^{2j-1} := \bar{v}^j \quad \text{and} \quad w^{2j} := \tfrac{1}{2}(\bar{v}^j + \bar{v}^{j+1}), \tag{9}$$

for $j = 1, 2, \ldots, m$. In fact, the values of coarse grid points are mapped unchanged to the fine grid and the values of the fine grid points which are not on the coarse grid, are equal to the average of the point values in their neighborhood of coarse grid. In this way, we obtain a vector of the form $w = (w^1, w^2, \ldots, w^{2m-1})^T$ which is corresponding to a fine grid. It should be noted that one may apply the operator I_k several times (e.g., r) to obtain the right Lanczos vector for the matrix A. The following algorithm describes the procedure clearly:

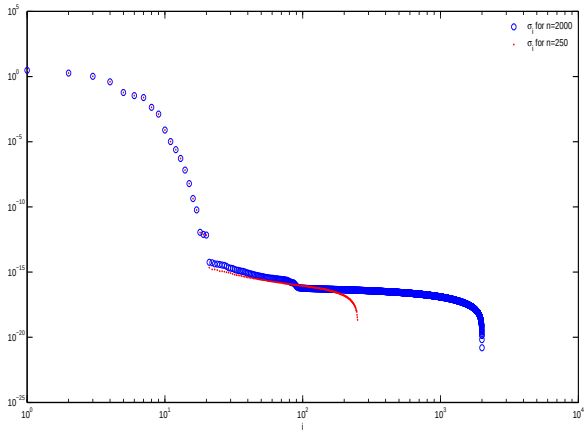

Figure 3. A comparison between singular values of *shaw* test problem for different dimensions of A.

Algorithm 2. Computing right Lanczos vectors of A from $\overline{v}$

1. Input $A = [a_1, \cdots, a_n]$, $\bar{A} = [\bar{a}_1, \cdots, \bar{a}_m]$, $b \in \mathbb{R}^n$, $\bar{b} \in \mathbb{R}^m$, $\overline{v} \in \mathbb{R}^m$ and $r \in \mathbb{N}$.
2. For $s = 1$ to r do
3. $\quad m := \text{length}(\overline{v})$
4. $\quad w := \text{zeros}(2m-1, 1)$
5. $\quad j := 1$
6. $\quad$ For $i = 1$ to m do
7. $\quad\quad w^j := \overline{v}^i$
8. $\quad\quad$ If $i \neq m$, then
9. $\quad\quad\quad w^{j+1} = (\overline{v}^i + \overline{v}^{i+1})/2$
10. $\quad\quad\quad j = j + 2$
11. $\quad\quad$ Endif
12. $\quad$ Endfor
13. $\quad \overline{v} := w$
14. EndFor
15. $q := 1$
16. If $\bar{a}_q^T \bar{b} \neq 0$ and $a_q^T b \neq 0$, then
17. $\quad \bar{v}_{q1} = \frac{\bar{a}_q^T \bar{b}}{\|\bar{A}^T \bar{b}\|}$,
18. $\quad v_{q1} = \frac{a_q^T b}{\|A^T b\|}$,
19. $\quad \alpha = \frac{\bar{v}_{q1}}{v_{q1}}$,
20. Else
21. $\quad q := q + 1$
22. $\quad$ Go to Line 16
23. EndIf
24. Return $\alpha w \in \mathbb{R}^{2^r(m-1)+1}$

Remark 1: We note that $2^r(m-1)+1 = n$ and so

$$m = (n-1)2^{-r} + 1. \tag{10}$$

Hence, in discretization of ill-posed problems like Fredholm equation of the first kind, one can choose n such that $(n-1)2^{-r} + 1$ is a natural number for some small r (e.g $r < 10$).

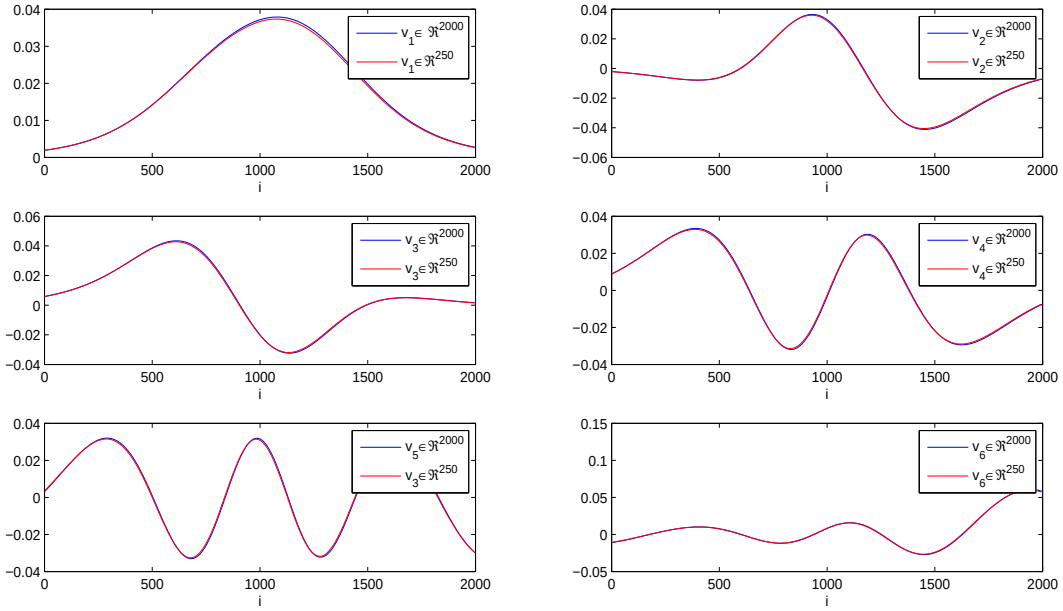

Figure 4. A comparison between right Lanczos vectors of *shaw* test problem for different dimensions of A.

Convergence of Algorithm 2

In order to show the convergence of Algorithm 2, we take simply for $i = 1, 2, \ldots, r$, the vectors u_i and v_i to be the exact and approximated right Lanczos vectors of the discretized system on the i-th coarse grid, respectively. In addition, let I_{i+1}^i be the linear interpolation operator such that it takes vectors v_{i+1} in the $(i+1)$-th coarse grid and produces vectors v_i in the i-th fine-grid according to the rule $I_{i+1}^i v_{i+1} = v_i$. Now, we assume that

$$\|u_i - I_{i+1}^i u_{i+1}\| \leq \epsilon_i, \tag{11}$$

which ϵ_i denotes the accuracy of the approximation. We have:

$$v_i = I_{i+1}^i I_{i+2}^{i+1} \ldots I_r^{r-1} v_r, \quad i = 1, 2, \ldots, r. \tag{12}$$

On the other hand, because of the boundedness of the interpolation operator, there exists a constant α_i such that for any vectors v_i and w_i in the i-th grid,

$$\|I_{i+1}^i v_{i+1} - I_{i+1}^i w_{i+1}\| = \|I_{i+1}^i (v_{i+1} - w_{i+1})\| \leq \alpha_i \|v_{i+1} - w_{i+1}\|, \tag{13}$$

holds. Then, by (11), (12) and (13), we have:

$$\|u_1 - v_1\| = \|u_1 - I_2^1 v_2\| \leq \|u_1 - I_2^1 u_2\| + \|I_2^1 u_2 - I_2^1 v_2\| \leq \epsilon_1 + \alpha_1 \|u_2 - v_2\|.$$

Similarly, we have:

$$\|u_i - v_i\| \leq \epsilon_i + \alpha_i \|u_{i+1} - v_{i+1}\|, \quad i = 1, 2, \ldots, r-1.$$

Hence,

$$\|u_1 - v_1\| \leq \epsilon_1 + \alpha_1(\epsilon_2 + \alpha_2(\ldots(\epsilon_{r-1} + \alpha_{r-1}\|u_r - v_r\|)\ldots)).$$

For the coarsest level r, we assume that $\|u_r - v_r\| \leq \epsilon_r$. So, $\epsilon = \epsilon_1 + \alpha_1(\epsilon_2 + \alpha_2(\ldots(\epsilon_{r-1} + \alpha_{r-1}\epsilon_r)\ldots))$ is an upper bound for $\|u_1 - v_1\|$. Now, if $\epsilon_i \to 0$ for $i = 1, 2, \ldots, r$, then $\|u_1 - v_1\| \to 0$ and convergence is guaranteed.

Remark 2: We note that

$$\|u_1 - v_1\| = \|u_1 - I_2^1 I_3^2 \dots I_r^{r-1} v_r\|, \tag{14}$$

where $v_r = \overline{v}$. Then, by (14) we can measure the difference between the exact right Lanczos vector u_1 and that of approximated vector obtained using $\overline{v}$ in Algorithm 2 for different values of m.

Algorithm 3. `Left-Preconditioned` CGLS $M_{MG}A^TAx = M_{MG}A^Tb$

1. `Let` $r^{(0)} = b - Ax^{(0)}, \quad p^{(0)} = s^{(0)} = M_{MG}(A^Tr^{(0)}), \quad \gamma_0 = \| s^{(0)} \|_2^2$
2. `For` $j = o$ `until convergence do`
3. $\quad t^{(j)} = M_{MG}p^{(j)}$,
4. $\quad q^{(j)} = At^{(j)}$,
5. $\quad \alpha^{(j)} = \frac{\gamma_j}{\|q^{(j)}\|_2^2}$,
6. $\quad x^{(j+1)} = x^{(j)} + \alpha_j t^{(j)}$,
7. $\quad r^{(j+1)} = r^{(j)} - \alpha_j q^{(j)}$,
8. $\quad s^{(j+1)} = M_{MG}(A^Tr^{(j+1)})$,
9. $\quad \gamma_{j+1} = \| s^{(j+1)} \|_2^2$,
10. $\quad \beta_j = \frac{\gamma_{j+1}}{\gamma_j}$,
11. $\quad p^{(j+1)} = s^{(j+1)} + \beta_k p^{(j)}$,
12. `Endfor`

Let $\bar{V}_k = [\bar{v}_1, \cdots, \bar{v}_k]$ with $\bar{v}_i,\ i = 1, \cdots, k$, as the right Lanczos vectors for the coarsest grid. We define $I_k^r\bar{V}_k = [I_k^r\bar{v}_1, \cdots, I_k^r\bar{v}_k]$, where I_k^r is the prolongation operator and it is applied r-times in Algorithm 2. In fact, by Algorithm 2, $I_k^r\bar{v} = w$. Now, the new form of the preconditioner can be defined as

$$M_{MG} = \alpha^2(I_k^r\bar{V}_k)(B_k^TB_k)^{-1}(I_k^r\bar{V}_k)^T + (I - \alpha^2(I_k^r\bar{V}_k)(I_k^r\bar{V}_k)^T).$$

We call this new preconditioner by subscript MG, because this preconditioner is based on an idea of prolongation and restriction operators in Multigrid algorithms. The construction of regularized inverse preconditioners M and M_{MG} need $O(kn^2)$ and $O(km^2 + rm)$ operations, respectively. We know, k and r are very small in comparison with m and n. On the other hand, by (10), the value of m is much smaller than n. Therefore, computational cost of the new form Lanczos based preconditioner M_{MG} for large n is more less than that of the preconditioner M. This preconditioner can be used as a left or right preconditioner with the CGLS method [14, 15]. Algorithm 3 shows the scheme of a left preconditioned CGLS method. In fact, CGLS is a semiconvergent method: For some j, in the first j iterations, the method converges to the exact solution, and then suddenly starts to diverge and the noise begins to enter the solution. Methods for finding the optimal value of such j (e.g. L-curve) can be found in [7, 9].

To illustrate the efficiency of the proposed method, we consider the *shaw* test problem for two cases of dimensions 250 and 2000. Figure 3 shows the singular values obtained by SVD from the discretized matrix of *shaw* problem, with the dimensions of 2000 and 250 for the matrices A and $\bar{A}$, respectively. As we observe, the value and number of large singular values of A and $\bar{A}$ are almost the same. Also, Figure 4 shows a comparison between right Lanczos vectors of *shaw* problem for different dimensions of A with the modifier coefficient $\alpha = \frac{v_{11}}{\bar{v}_{11}} \approx \frac{1}{3}$.

5. Numerical results

In this section, we present some numerical examples. All experiments were carried out using MATLAB and Hansen Regularization [8]. We investigate the performance of our proposed regularization parameter (k) for introduced preconditioner and LSQR method. In the following examples, taken from [8] the contaminated vector b and the Gaussian noise e are considered as $b = \hat{b} + e$, where

$$e = 10^{-3}\|\hat{b}\|.\frac{\omega}{\|\omega\|},$$

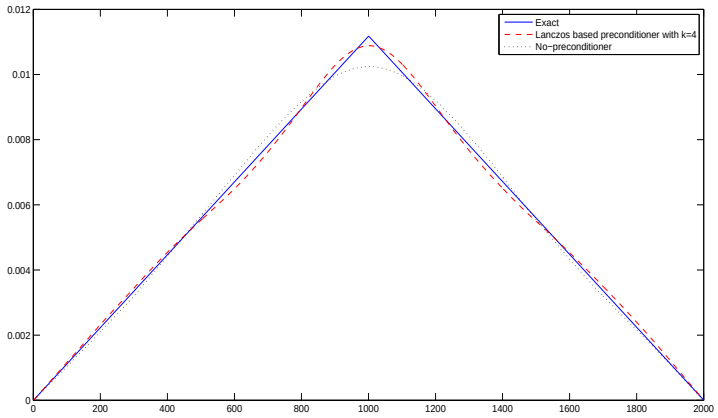

Figure 5. A comparison between the exact solution and the solution obtained by preconditioned CGLS method at step 2 and the selection $k = 4$ of bidiagonalization Lanczos for **derive2(2000,3)** problem.

and ω is Gaussian random vector with mean 0 and standard deviation 1. In all the following examples, we consider the coarse grid with dimensions 250×250. On this coarse grid, the parameter k is obtained by (7). Moreover, by using the interpolation operators, the ingredients of the preconditioner M, i.e. V_k, σ_k and B_k are obtained for fine grid.

Example 5.1 Consider the integral equation

$$\int_0^1 K(s,t)f(t)dt = g(s), \qquad 0 \le s \le 1,$$

where

$$K(s,t) = \begin{cases} s(t-1), & s < t, \\ t(s-1), & s \ge t, \end{cases}$$

$$f(t) = \begin{cases} t, & t < \frac{1}{2}, \\ 1-t, & t \ge \frac{1}{2}, \end{cases}$$

$$g(s) = \begin{cases} \frac{(4s^3-3s)}{24}, & s < \frac{1}{2}, \\ \frac{(-4s^3+12s^2-9s+1)}{24}, & s \ge \frac{1}{2}. \end{cases}$$

This is a mildly ill-posed problem; i.e., its singular values decay slowly to zero. We discretize this integral equation, using the function **derive2(2000,3)** (see [8]), to obtain the matrix $A \in \mathbb{R}^{2000\times 2000}$.

We apply the introduced methods in Section 4 for selected suitable regularization parameter on this coarse grid. In this problem, we have only used $k = 4$ steps of Lanczos bidiagonalization for constructing the preconditioner. Figure 5 shows the solution of Lanczos preconditioned obtained for $k = 4$ steps of the bi-diagonalization Lanczos and unpreconditioned systems after two iteration steps. Also, Figure 6 shows that the relative error for preconditioned linear systems with selection $k = 4$ is better than that of the original system.

Example 5.2 The inverse heat equation used here is a Volterra integral equation of the first kind with $[0, 1]$ as integration interval. The kernel is $K(s,t) = K(s-t)$ with

$$k(t) = \frac{t^{\frac{-3}{2}}}{2\mathbf{kappa}\sqrt{\pi}} \exp(-\frac{1}{4\mathbf{kappa}^2 t}).$$

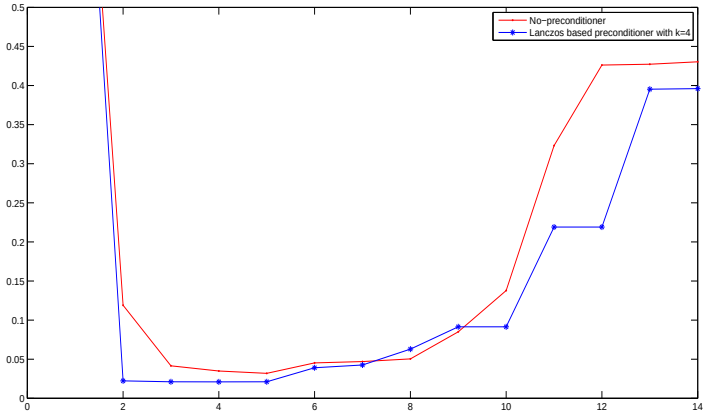

Figure 6. Relative errors versus iteration steps for $A^TAx = A^Tb$ of **derive2(2000,3)** problem

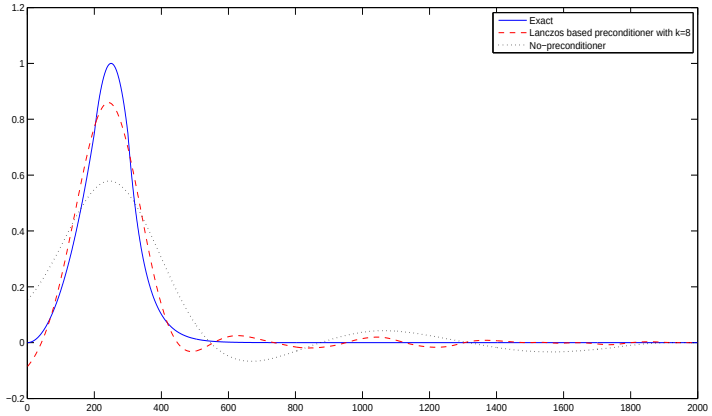

Figure 7. A comparison between the exact solution and the solution obtained by preconditioned CGLS method with 4 iterations and the selection $k = 8$ of Lanczos bidiagonalization process for $A^TAx = A^Tb$ of **heat** problem.

Here, the parameter **kappa** controls the ill-conditioning of the matrix A:
kappa =5 gives a well-conditioned matrix,
kappa =1 gives an ill-conditioned matrix.
The default is **kappa**=1.
The integral equation is discretized by means of simple collocation and the midpoint rule with n grid points (see [5, 6]). An exact solution x is constructed, and then the right-hand side b is produced as $b = Ax$. We use the code **heat(n,kappa)** from [8] to discretize this integral to obtain the matrix $A \in \mathbb{R}^{2000\times 2000}$.

In this problem, we used the regularization parameter $k = 8$. Figure 7 compares the solution obtained of preconditioned system at 4 iterations with $k = 8$ steps of Lanczos bidiagonalization and without preconditioner. Figure 8 shows the relative error with and without preconditioner.

Example 5.3 We discretize a Fredholm integral equation of the first kind (2) with $[-\frac{\pi}{2}, \frac{\pi}{2}]$ as both integration interval. The kernel K and solution f are given by

$$K(s,t) = (\cos s + \cos t)^2(\frac{\sin u}{u})^2, \qquad u = \pi(\sin s + \sin t),$$

$$f(t) = a_1 \exp(-c_1(t-t_1)^2) + a_2 \exp(-c_2(t-t_2)^2).$$

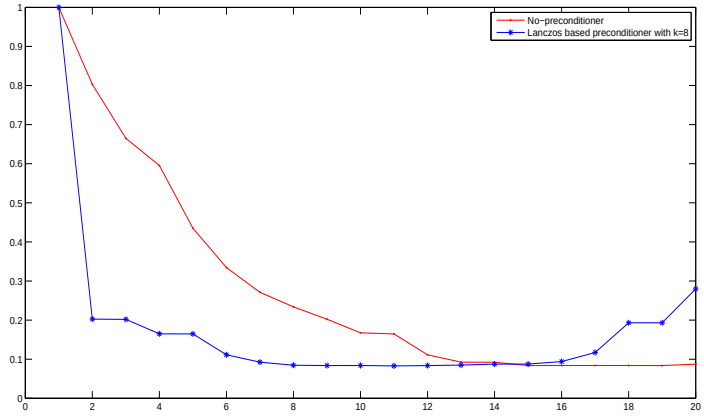

Figure 8. Relative errors versus iteration steps, for $A^TAx = A^Tb$ of **heat** problem.

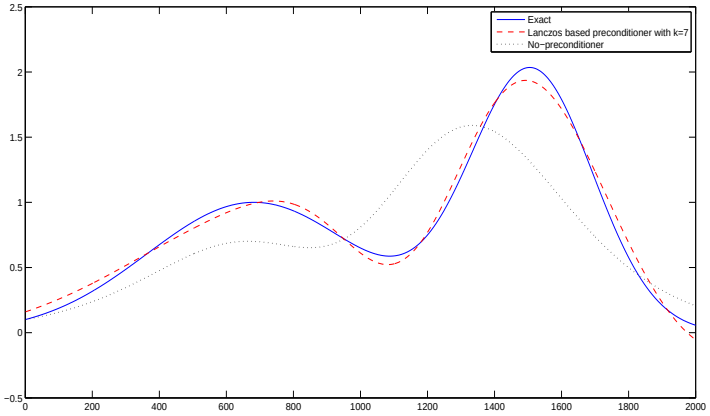

Figure 9. A comparison between the exact solution and the solution obtained by preconditioned $CGLS$ method with 3 iterations and the selection $k = 7$ of bidiagonalization Lanczos for $A^TAx = A^Tb$ of **shaw** problem.

The parameters a_1, a_2, c_1, c_2, t_1 and t_2 are constants that determine the shape of the solution f. In this implementation, we use $a_1 = 2, a_2 = 1, c_1 = 6, c_2 = 2, t_1 = 0.8$ and $t_2 = -0.5$. The kernel and the solution are discretized by simple collocation with n points to produce A and x. Then the discrete right-hand side is produced by $b = Ax$. We used the code **shaw** from [7] to obtain $A \in \mathbb{R}^{2000\times 2000}$.

In Figure 9, we plotted the solution obtained by Lanczos based preconditioner with $k = 7$ at step 2 and unpreconditioned systems. Figure 10 shows, the relative errors from iteration steps of CGLS method for preconditioned systems with $k = 7$ and original system.

Example 5.4 Define the function

$$\phi(x) = \begin{cases} 1 + \cos\frac{\pi x}{3}, & |x| < 3, \\ 0, & |x| \geq 3. \end{cases}$$

Then kernel K, the solution f, and the right-hand side g are given by

$$\begin{aligned} K(s,t) &= \phi(s-t), \\ f(t) &= \phi(t), \\ g(s) &= (6-|s|)(1+\frac{1}{2}\cos\frac{\pi x}{3}) + \frac{9}{2\pi}\sin\frac{\pi|s|}{3}. \end{aligned}$$

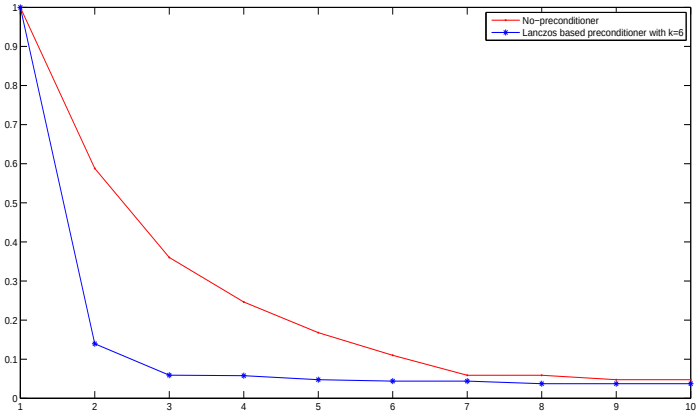

Figure 10. Relative errors versus iteration steps, for $A^TAx = A^Tb$ of **shaw** problem.

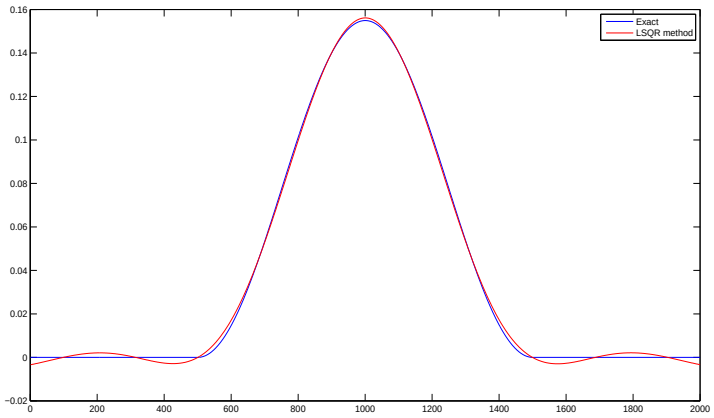

Figure 11. A comparison between the exact solution and the solution obtained by LSQR method at $k = 6$ for $\min \parallel b - Ax \parallel$ of **phillips** problem.

Both integration intervals are $[-6, 6]$. We discretize this problem, using the function **phillips(2000)**[12], to obtain the matrix $A \in \mathbb{R}^{2000\times 2000}$. In the example, we used (6) to compute the approximated solution. Figure 11, displays an exact solution and the solution obtained by LSQR method for $k = 6$ steps of Lanczos bidiagonalization.

6. Conclusion

We have studied a recently proposed preconditioner for ill-posed linear system of equations which is based on the Lanczos bidiagonalization method. Then, we have presented an efficient method to produced such preconditioner. Finally, some numerical results have been presented to show the effectiveness of the proposed method.

References

[1] R.C. Allen, W.R. Boland, V. Faber, G.M. Wing, Singular values and condition numbers of Galerkin matrices arising from linear integral equations of the first kind, J. Math. Anal. Appl. 109(1985) 564-590.
[2] A. Bjorck, Numerical methods for least square problems, SIAM, Philadelphia, 1996.
[3] S.C. Brenner, L.R. Scott, The mathematical theory of finite element methods, 3rd edition, Springer, 2008.
[4] D. Calvetti, G. Golub, L. Reichel, Estimation of the L-curve via Lanczos bidiagonalization, BIT. 39(1999) 603-619.
[5] A.S. Carasso, Determining surface temperatures from interior observations, SIAM J. Appl. Math. 42(1982) 558-574.
[6] L. Elden, The numerical solution of a non-characteristic Cauchy problem for a parabolic equation, in: Deuflhart, P., Hairer, E. (Eds.) Numerical Treatment of Inverse Problems in Differential and Integral Equations, Birkhauser, Boston, 1983.

[7] P.C. Hansen, Rank-deficient and discrete ill-posed problems, SIAM, Philadelphia, 1998.
[8] P.C. Hansen, Regularization tools: a MATLAB package for analysiss and solution of discrete ill-posed problems, Numer. Alg. 6(1994) 1-35.
[9] P.C. Hansen, T.K. Jensen, G. Rodriguez, An adaptive pruning algorithm for the discrete L-curve criterion, J. Comput. Appl. Math. 198(2007) 483-492.
[10] P.C. Hansen, Computation of the Singular Value Expansion, Computing 40(1998) 185-199.
[11] F.R. de Hoog, Review of Fredholm equations of the first kind, The application and numerical solution of integral equations.Sijthoff and Noordhoff, pp. 119-134. Leyden, Netherlands, 1980.
[12] M.E. Kilmer, D.P. Oleary, Choosing regularization parameter in iterative methods for ill-posed problems, SIAM J. Matr. Anal. Appl. 22(2001) 1204-1221.
[13] A. Kleefeld, Numerical results of linear Freadholm integral equations of the first kind over surfaces, University of Wisconsin-Milwaukee, 2007.
[14] D.P. Oleary, J.A. Simmons, A bidiagonalization-regularization procedure for large scale discretizations of ill-posed problems, SIAM J. Sci. and Stat. Comput. 2(1981) 474-489.
[15] M. Rezaghi, S.M. Hosseini, Lanczos based preconditioner for discrete ill-posed problems. Computing. 88(2010) 79-76.
[16] Y. Saad, Iterative Methods for Sparse linear Systems, PWS press, New York, 1995.
[17] F. Smithies, The eigenvalues and singular values of integral equations, Proc. London Math. Soc. 43(1937), 255-279.