A Preconditioned GMRES Method for Solving a 1D Sideways Heat Equation

Zohreh Ranjbar and Lars Eldén

LiTH-MAT-R, 0348-2960, No. 6

Available at: Linköping University Electronic Press
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-61152
A Preconditioned GMRES Method for Solving a 1D Sideways Heat Equation

Zohreh Ranjbar* and Lars Eldén*
Department of Mathematics, Linköping University
SE-581 83 Linköping, Sweden.

November 9, 2010

Abstract

The sideways Heat equation (SHE) is a model of the problem of determining the temperature on the surface of a body from the interior measurements. Mathematically it can be formulated as a non-characteristic Cauchy problem for a parabolic partial differential equation. This problem is severely ill-posed: the solution does not depend continuously on the data. We use a preconditioned Generalized Minimum Residuals Method (GMRES) to solve a 1D SHE. Regularization is used in the preconditioner as well as truncating the GMRES algorithm. Numerical experiments demonstrate that the proposed method works well.

Keywords: Cauchy problem, inverse problem, ill-posed, iterative methods, GMRES preconditioning, FFT, heat equation, truncated SVD

1 Introduction

In several dynamical heat transfer situations [1, 5] it is required to determine the surface temperature of a body, where the surface itself is inaccessible to measurements. Mathematically this can be formulated as a Cauchy problem for a parabolic equation, which we will call a sideways heat equation (SHE)\(^1\). This type of problem is known to be severely ill-posed: A small perturbation

---

*E-mail: zohreh.ranjbar@liu.se, lars.elden@liu.se.

\(^1\)In some literature [1] it is called the Inverse Heat Conduction Problem (IHCP).
of the data can induce a large error in the solution. Due to ill-posedness and unavoidable noise (measurement errors) in the data we must use a regularization method to compute a stabilized solution that is less sensitive to the noise. By regularization the noise is filtered out at the cost of restricting the set of admissible solutions.

There are various regularization methods that could be used for solving the SHE, e.g., Tikhonov regularization, truncated SVD, and iterative methods, see [13] for an overview. With two- and three-dimensional problems with variable coefficients in mind, we propose to solve the SHE using an iterative method, preconditioned GMRES (generalized minimum residual [21]), which is a Krylov subspace method. The regularizing properties of the GMRES method have recently been studied in several papers [17, 2, 3, 4, 14, 16], and also preconditioned GMRES has been proposed [16].

We consider a SHE in one space dimension with constant coefficient \( a > 0 \),

\[
\begin{align*}
  u_t &= au_{xx}, & 0 < x < 1, & 0 \leq t \leq 1, \\
  u(x, 0) &= 0, & 0 \leq x \leq 1, \\
  u(1, t) &= g(t), & 0 \leq t \leq 1, \\
  u_x(1, t) &= 0, & 0 \leq t \leq 1.
\end{align*}
\]

(1.1)

We wish to determine the value at the left boundary, \( u(0, t) = f(t) \), from the data at the right boundary \( u(1, t) = g(t) \). After discretization the problem can be formulated as a least squares problem,

\[
\min_f \| Kf - g \|_2 \quad \text{(1.2)}
\]

where \( K \) is a lower triangular Toeplitz matrix, \( K \in \mathbb{R}^{n \times n} \), and \( f \) and \( g \) are vectors. The matrix \( K \) is extremely ill-conditioned for large enough dimension \( n \). We use a Chan-type [7] circulant preconditioner to precondition the discrete system (1.2). Since the preconditioner \( C \) is computed from the ill-conditioned matrix \( K \), we can expect that it too is ill-conditioned. Therefore, we truncate its eigenvalues so that only those that are large enough are retained. We demonstrate that for this one-dimensional problem the GMRES iteration, with truncated circulant preconditioner, gives a good approximate solution in very few iterations. The accuracy of the solution is shown to be comparable to that of standard methods, like Tikhonov regularization.

When the coefficient \( a \) is not constant we have no explicit representation of \( K \). Therefore we propose to solve it numerically using preconditioned GMRES, where each application of the operator corresponding to the matrix \( K \) in (1.2) is performed by solving a well-posed parabolic problem. As preconditioner we use the solution of a 1D SHE with constant coefficients, expressed in terms of a truncated singular value decomposition.
The outline of the paper is as follows. In Section 2, we give some background knowledge about the preconditioned GMRES (PGMRES) method and discuss the effect of preconditioning of ill-posed problems. The solution of the 1D SHE with both constant and variable coefficient using preconditioned GMRES is presented in Section 3. We provide some numerical results to illustrate that our schemes work well. Concluding remarks are given in section 4.

Throughout the paper the notation $\| \cdot \|$ denotes the $L^2$-norm. When we are dealing with vectors we use the Euclidean norm, denoted $\| \cdot \|_2$.

## 2 Preconditioned GMRES

Krylov subspace methods are the methods of choice for the iterative solution of large and sparse linear systems $Ax = b$. The idea is to find an approximate solution $x_m$ in the Krylov subspace

$$K_m(A, r_0) = \text{span}\{r_0, Ar_0, A^2r_0, \ldots, A^{m-1}r_0\},$$

where $r_0 = b - Ax_0$ is the initial residual and $x_0$ is an initial guess of the solution. The Generalized Minimum Residual Method (GMRES) by Saad and Schultz [21] is the standard Krylov subspace method for problems with non-symmetric matrix $A$. GMRES generates a sequence of orthonormal vectors $v_1, v_2, \ldots$, such that $\beta v_1 = r_0$ and $v_1, \ldots, v_m$ is a basis in $K_m(A, r_0)$, see [20, Section 6.5]. Furthermore,

$$AV_m = V_mH_m + h_{j+1,j}v_{m+1}^Te_m, \quad V_m = [v_1, \ldots, v_m],$$

where $H_m = (h_{ij}) \in \mathbb{R}^{m \times m}$ is an upper Hessenberg matrix, and $h_{j+1,j} = \|Av_j - \sum_{i=1}^j h_{ij}v_i\|$. It can be seen that the GMRES approximation is the vector in $x_0 + K_m$, that minimizes $\|b - A(x_0 + V_my)\|$. Equivalently,

$$x_m = x_0 + V_my_m, \quad y_m = \text{argmin}_y \|\beta e_1 - H_my\|.$$

In the iterative solution of well-posed problems using a Krylov subspace method one usually precondition the problem, e.g., in the case of a right preconditioner, by considering the equivalent system of equations

$$AM^{-1}z = b, \quad z = Mx,$$

where $M$ is an approximation of $A$, such that systems $Mx = z$ can easily be solved. In the case of symmetric $A$, the effectiveness of a preconditioner is determined by its ability to cluster the eigenvalues of the preconditioned matrix.
around a point away from the origin. Clustering the eigenvalues increases the rate of convergence of iterative methods. In the non-symmetric case there is no immediate relation between the rate of convergence of the preconditioned Krylov method and the clustering of eigenvalues. Still, preconditioning with a matrix that is a good approximation of $A$ often works very well. Moreover, it is of interest to choose a $M$ such that the vector $M^{-1}z$ can be evaluated rapidly. Since in GMRES we are computing an approximate solution by solving a least squares problem, see the algorithm description below, it is the approximation and clustering of singular values that give insight into the rate of convergence of the Krylov method.

The preconditioned GMRES (PGMRES) algorithm described below is given in [20].

**GMRES with Right Preconditioning**

1. Compute $r_0 = b - Ax_0$, $\beta = \|r_0\|_2$, and $v_1 = r_0/\beta$.
2. for $j = 1, \ldots, m$,
   Compute $w := AM^{-1}v_j$
   for $i = 1, \ldots, j$ do
     $h_{i,j} := w^Tv_i$
     $w := w - h_{i,j}v_i$
   end
   Compute $h_{j+1,j} = \|w\|_2$ and $v_{j+1} = w/h_{j+1,j}$
3. Define $V_m := [v_1, \ldots, v_m]$, $H_m = \{h_{i,j}\}_{1 \leq i \leq j+1, 1 \leq j \leq m}$
4. Compute $y_m = \arg\min_y \|\beta e_1 - H_m y\|_2$ and $x_m = x_0 + M^{-1}V_m y_m$

A preconditioned Krylov method generates a Krylov subspace

$$\mathcal{K}_m(\tilde{A}, r_0) = \text{span}\{r_0, \tilde{A}r_0, \tilde{A}^2r_0, \ldots, \tilde{A}^{m-1}r_0\}, \quad \tilde{A} = AM^{-1}.$$  

It is well-known that Krylov space methods give good solutions of ill-posed problems already after a small number of steps, and that after a certain number of steps the error in the computed solution starts to grow, so called “semi-convergence” behavior cf. [19, p. 259]. This is due to the fact that in its first few steps the Krylov method approximates mainly solution components associated with the largest singular values. Then gradually the smaller singular values start to influence the solution, and finally, after many steps, those singular values that are close to zero are approximated, which makes the solution explode.
Thus clustering of all the singular values (eigenvalues) can not be the objective of preconditioning of ill-posed problems, since we are only interested in approximate solutions that are spanned by the singular vectors corresponding to the largest singular values. Instead we want to find a preconditioner that approximates the largest singular values and the corresponding singular vectors well, and which suppresses the influence of the small singular values. We describe such a preconditioner in the following section.

3 The One-Dimensional Sideways Heat Equation

Here we consider a one-dimensional SHE,

\[ u_t = au_{xx}, \quad 0 < x < 1, \quad 0 \leq t \leq 1, \]
\[ u(x, 0) = 0, \quad 0 \leq x \leq 1, \]
\[ u(1, t) = g(t), \quad 0 \leq t \leq 1, \]
\[ u_x(1, t) = 0, \quad 0 \leq t \leq 1. \]  

(3.1)

where the value at the left boundary, \( u(0, t) = f(t) \), is sought. For concreteness we assume \( a = 1 \). Another formulation of the Cauchy problem (1.1) [9] is given by

\[ g(t) = \int_0^t k(t - \tau)f(\tau) \, d\tau, \quad 0 \leq t \leq 1, \]  

(3.2)

where the kernel function is defined by

\[ k(t) = 2 \sum_{n=0}^{\infty} (-1)^n \nu_n \exp(-\nu_n^2 t), \quad \nu_n = (2n + 1)\pi/2. \]

Discretization of the Volterra integral equation (3.2) leads to a linear system of equations,

\[ Kf = g, \]

(3.3)

(or a least squares problem (1.2)), where \( K \in \mathbb{R}^{n \times n} \) is a lower triangular Toeplitz matrix,

\[ K = \begin{pmatrix} k_0 & k_0 & \cdots & \cdots & \cdots \\ k_1 & k_0 & \cdots & \cdots & \cdots \\ \vdots & \vdots & \ddots & \ddots & \cdots \\ k_{n-1} & \cdots & k_1 & k_0 \end{pmatrix}, \]

(3.4)

and \( f \) and \( g \) are vectors.
The matrix $K$ is extremely ill-conditioned for large enough dimension $n$. In our illustrations below the dimension $n$ is chosen equal to 200. Due to ill-conditioning and unavoidable noise (measurement errors) in the data vector $g$, one cannot expect to compute an accurate approximation of $f$ by simply solve the linear system (3.3). We want to solve the linear system by PGMRES and regularization will be incorporated in the preconditioner.

### 3.1 Circulant Preconditioning

The main purpose of using a preconditioner in the iterative solution of a linear systems of equations is to reduce the number of iterations. In the context of discrete ill-posed problems preconditioners may also improve the quality of the computed solution [14]. Circulant preconditioners for linear systems with Toeplitz structure have been studied for quite a long time, see e.g., [6] for a survey. In the context of an ill-posed problem we want the preconditioner to reduce the number of iterations. This can be done by clustering of singular values. However, since it is the small singular values that account for the ill-conditioning, we do not want those to be clustered along with the larger singular values. We propose to use the special structure of a circulant preconditioner in such a way that it is mainly the large singular values that are emphasized, and the smaller singular values are suppressed. We have chosen the preconditioner $C$ as Chan’s circulant preconditioner [7], which minimizes $\|K - C\|_F$ over all circulant matrices $C$. Recall that a matrix is called circulant if it is Toeplitz and the last entry of every row is the first entry of its succeeding row. A circulant matrix

$$
C = \begin{pmatrix}
    c_0 & c_{n-1} & \cdots & c_2 & c_1 \\
    c_1 & c_0 & c_{n-1} & \cdots & c_2 \\
    c_2 & c_1 & c_0 & \ddots & \vdots \\
    \vdots & \vdots & \ddots & \ddots & c_{n-1} \\
    c_{n-1} & \cdots & \cdots & c_1 & c_0, \\
\end{pmatrix}
$$

is completely specified by its first column, $c = (c_0 \ c_1 \ \cdots \ c_{n-1})^T$. It is well-known that every $n \times n$ circulant matrix $C$ can be diagonalized by fast Fourier transform

$$
C = F \Lambda F^*,
$$

where the entries of Fourier matrix $F$ are given by

$$
(F)_{jk} = \frac{1}{\sqrt{n}} \exp^{2\pi i j k / n}, \quad 0 \leq j, k \leq n - 1, \quad i = \sqrt{-1},
$$
see [22, p. 3], and
\[ \Lambda = \text{diag}\left(\lambda_1 \lambda_2 \cdots \lambda_n\right)^T = \text{diag}(F^*c). \]

The Fourier matrix is unitary. It is a standard result in matrix theory, see e.g. [8, p. 72], [22, p. 206], [6, p. 11]. For more details and properties of circulant matrices, see Davis[8]. It should be noted that neither the Fourier matrix nor the preconditioner need be explicitly formed. Instead, $F$ and $F^*$ are applied using FFT and $\Lambda$ is stored as a vector, which is computed by applying FFT on $c$ (the first column of $C$). The matrix-vector multiplications $Fy$ and $F^*y$ can be performed in $O(n \log n)$ operations by the Fast Fourier Transform (FFT).

This preconditioner $C$ is constructed by averaging the entries along pairs of diagonals of $K$. Here, due to the lower triangular structure of $K$ (3.4), the entries of $C$ are chosen as
\[ c_0 = k_0, \quad c_i = \frac{(n-i)k_i}{n}, \quad i = 1, \ldots, n-1. \]  

(3.6)

We illustrate the matrices $K$ and $C$ in Figure 3.1. We now will demonstrate that $C$ can be considered as an approximation of $K$ in the sense that its first singular values are quite close to those of $K$ (comparing eigenvalues does not make sense, as $K$ has the eigenvalue $k_0$ of multiplicity $n$ and almost all eigenvalues of $C$ are complex). In Figure 3.2 we illustrate the singular values of $K$ and $C$. It is seen that the first 100 singular values of $C$ are quite good approximations of those of $K$. Obviously, since $K$ is very ill-conditioned, we do not want to approximate all the singular values of $K$ using the preconditioner, only the largest ones. Therefore we choose the preconditioner as a truncated eigenvalue expansion of $C^{-1}$:
\[ KC^†_Tz = KF^*A^†_T Fz, \]

where
\[ (A^†_T)_{jj} = \begin{cases} 1/\lambda_j & |\lambda_j| > \tau, \\ 0 & |\lambda_j| \leq \tau. \end{cases} \]  

(3.7)

Determining the a suitable value of $\tau$ is closely related to the problem of selecting a regularization parameter. If it is chosen too small, then we may remove smooth solution components that are relevant. On the other hand, if it is chosen too large, then very little preconditioning is done, and the convergence is not accelerated as was the aim. A method for choosing $\tau$ [12] is described below. Also recall that in an iterative regularization method the number of iterations performed constitutes a regularization parameter.
Figure 3.1: The matrices $K$ (top) and $C$ (bottom).

Figure 3.2: All singular values of $K$ (left). The first 195 singular values of $K$ (o) and all those of $C$ (×) (right).
Therefore the termination criterion is also important. The “condition L-curve” has been used for determining when to terminate the iteration, see [3]. This method is based on the condition number of projected matrices defined by the GMRES method. Actually in classic L-curve method [15] we try to find the vertex on the piecewise linear graph that connects adjacent points in the sequence \((\|x_k\|, \|r_k\|)\) for \(k = 1, 2, \ldots\), where \(x_k\) and \(r_k\) denote respectively the sequence of approximated solution and the associated residual vectors. But in the condition L-curve, \(\|x_k\|\) is replaced by \(\kappa(H_k)\), the condition number of the Hessenberg matrix \(H_k\) determined in the \(k\)th step.

### 3.2 Numerical Experiments

In this section we present two numerical examples. The first one is the one-dimensional SHE with constant coefficient \(a = 1\). The second example is also one-dimensional but with variable coefficient \(a(x)\).

**Example 1: Circulant preconditioner** We will now describe a numerical experiment where we compare GMRES with the truncated circulant preconditioner, denoted PGMRES, to an iterative least square solver based on Lanczos-Golub-Kahan bidiagonalization (LGK) [11], with and without preconditioning. We remark that GMRES without preconditioning did not work at all, cf. [16]. We selected a solution vector \(f\) and constructed the “exact” right hand side vector \(g\) by multiplying \(g = Kf\). Then the right hand side vector was perturbed, \(g_\delta = g + \delta\), by adding a “measurement error” vector, consisting of samples from a normal distribution (using the Matlab random number generator `randn`) and scaled such that

\[
\frac{\|g - g_\delta\|}{\|g\|} = 0.05.
\]

The data functions \(g\) and \(g_\delta\) are illustrated in Figure 3.3. The circulant matrix \(C\) was computed by (3.6). In order to find the truncation parameter \(\tau\) we used the method in [12, p. 169] which is based on the L-curve criterion [15, 18]. Thus, one selects the value of \(\tau\) that corresponds to the corner of L-curve for solving \(Cz = g_\delta\).² Note that it is essential that the data vector with noise is used, because otherwise there is usually no corner of the L-curve. In Figure 3.4 we illustrate the singular values of the matrices \(K\) and \(KC_\tau^\dagger\). In order to check the quality of the solution obtained using PGMRES we also give results for the case when there is no perturbation of the data. The results

²In [12] the eigenvalue expansion of the circulant matrix is not truncated, but the small eigenvalues are replaced by 1.
Figure 3.3: Data function without noise (left) and with 5% noise (right).

Figure 3.4: The first 180 singular values of $K^\ast$ and $KC^\dagger$, when the noise in the data function is zero (left) and 5% (right).

are quite good. In that case we adjusted the cut-off level manually so that the solution became as good as possible. We see in Figure 3.4 that the large singular values of the preconditioned matrix are clustered at one, and are well separated from the small singular values. One can also see the relation between the noise level and the truncation level: the L-curve criterion for $Cz = g$ gives a smaller value of $\tau$ than our manual choice for unperturbed data. In Figure 3.5 we compare the solution using PGMRES and LGK to Tikhonov regularization with smoothing matrix equal to the identity. The rate of convergence of PGMRES and LGK is illustrated in Figure 3.6, which displays the residual error $\|Kf^{(k)} - g\|$ as a function of iteration index $k$, and where $f^{(k)}$ denotes an iterate with either method. Finally, in Figure 3.7, we illustrate the relative error as a function of iteration index. It should be noted that in our experiments this preconditioning did not seem to work well together with LGK, as the relative errors of the solution were initially much larger than for the other two methods. In our experiments above using preconditioned GMRES the approximate solution was obtained with
Figure 3.5: Condition L-curve plots (left) for PGMRES method. Exact solution, and approximate solutions using PGMRES, LGK with and without preconditioning and Tikhonov regularization (right). The accuracies of the different approximate methods are comparable. The top plots are for noise-free data and the approximate solution computed after 12 iterations of PGMRES, 15 iterations LGK and 15 iterations PLGK. The bottom plot is for 5% perturbation and the approximate solution computed after 3 iterations of PGMRES, 8 iterations of LGK and 8 iterations of PLGK. The regularization parameter for Tikhonov method is chosen equal to $\tau$, the cut-off parameter in (3.7).
Figure 3.6: Residuals of PGMRES (solid with +) and LGK (solid with ◦) and preconditioned LGK (solid with ×) as functions of iteration index. The left plot is for noise-free data and the right plot for 5% perturbation.

Figure 3.7: Relative errors in the solution during 15 iterations of PGMRES (solid with +) and LGK (solid with *) and preconditioned LGK (solid with ×). The left plot is for noise-free data and the right plot for 5% perturbation.
fewer iteration than LGK method, although the rate of decay of the residuals were comparable. In addition, Figure 3.7 indicates that the solution is less sensitive to the actual number of iteration steps. Of course, this is due to the fact that the truncated preconditioner serves as a regularization method in itself. Our limited experiments also indicate that the accuracy of PGMRES is comparable to that of Tikhonov regularization.

**Example 2: Experiment with variable coefficient**

We here choose a SHE problem (3.1) with variable coefficient $a(x)$. In this case the kernel function of the operator $K$ in (3.3) is not known. Instead of using a circulant $C$ we assume that the preconditioner $M$ is a discretization of the corresponding equation with constant coefficient $\kappa$ equal to the mean value of $a(x)$ for $0 < x < 1$. Then $M$ is a Toeplitz matrix. Since the dimension of this problem is not very large, and since we are not at all concerned with the computational efficiency of the procedure, we apply the truncated singular value decomposition (TSVD) method to filter out the high frequency part of matrix $M^{-1}$. The TSVD solution of a linear system $Mx = b$ is computed as

$$x_p = \sum_{i=1}^{p} \frac{u_i^T b}{\sigma_i} v_i,$$

where

$$M = U \left( \begin{array}{c} \Sigma \\ 0 \end{array} \right) V^T,$$

is the SVD, and the singular vectors are denoted $u_i$ and $v_i$, respectively, and the singular values are denoted $\sigma_i$. In the TSVD method the useful information about the solution is represented by the largest singular values and corresponding vectors; therefore we neglect the undesired SVD components corresponding to small singular values. This procedure is analogous to the one used with circulant preconditioning (3.7), where we omitted small eigenvalues.

Let $U_p = [u_1, u_2, \ldots, u_p]$ consist of the $p$ first singular vectors, and correspondingly for $V_p$, and let $\Sigma_p = \text{diag}[\sigma_1, \sigma_2, \ldots, \sigma_p]$ be the $p \times p$ diagonal matrix of the first $p$ singular values. In the PGMRES method we evaluate

$$KM_p^\dagger x = KV_p \Sigma_p^{-1} U_p^T x.$$

The truncation parameter $p$ plays the role of regularization parameter. In each PGMRES iteration first we apply the preconditioner $M_p^\dagger$,

$$v = M_p^\dagger x = V_p \Sigma_p^{-1} U_p^T x,$$
and then we apply \( w = Kv \), which is equivalent to solving a well-posed 1D parabolic equation. We use the method of lines: a finite difference scheme in the \( x \)-direction combined with the Matlab ODE solver \texttt{ode23s} in the \( t \)-direction.

A numerical example was constructed with discontinuous coefficient \( a(x) \), shown in Figure 3.8, where we also have plotted the data function with 5% noise (the noise was added as in the previous example). The stopping criterion for the iteration is the condition L-curve, the same as in previous example. The results are shown in Figure 3.9, where the L-curve and the approximate solution are plotted. The accuracy of the solution is comparable to, or slightly better than that obtained with the method of lines approach of [10], where the time derivative was approximated by a bounded operator using a “spectral derivative”. Finally, in Figure 3.10, we illustrate the residual and relative error as a function of iteration index.

4 Conclusions

In this paper we have presented a new iterative regularization technique for solving a 1D sideways heat equation using a preconditioned GMRES.

The preconditioning operator is singular, and a pseudo-inverse is used. To investigate the applicability of the singular preconditioner, we used a truncated circulant preconditioner to a 1D SHE with constant coefficient, and found that it works well.

For the 1D SHE with variable coefficient the preconditioner is formulated in terms of truncated singular value decomposition.

The computed examples indicate that the proposed PGMRES methods
Condition number of the Hessenberg matrix

Figure 3.9: Condition L-curve plot for TSVD-based PGMRES method (left). Exact solution, approximate solution after 9 iterations of TSVD-PGMRES (dashed), and the approximate solution using the method of lines (dash-dotted).

Figure 3.10: Residual (left) and relative error (right) during 15 iterations of TSVD-based PGMRES applied the 1D SHE with variable coefficient $a(x)$ with 5% perturbation of the data.
are well suited for the solution of 1D sideways heat problems. To further investigate the usefulness of the method, we are planning to apply it to realistic problems with measured data. Future work also includes extension to multidimensional problems with variable coefficients.

References


