Ill-conditioned simultaneous linear equations to which the regularized solution almost coincides with the exact one

Takemi SHIGETA

Professor, Laboratory of Applied Mathematics, Showa Pharmaceutical University

Abstract

Ill-conditioned simultaneous linear equations in two unknowns are considered. A small noise contained in the non-homogeneous term of the ill-conditioned equations retains the potential to affect sensitively on the accuracy of the solution. The truncated singular value decomposition can then be applied to the equations to obtain a regularized solution of which the change is sufficiently small for a small change of the non-homogeneous term. In this study, the non-homogeneous term of the equations such that the regularized solution almost coincides with the exact one is defined to be exact. The purpose of the study is to reconstruct the unknown exact non-homogeneous term from the regularized solution. Simple numerical experiments show that the exact solution to the equations with the reconstructed non-homogeneous term almost coincides with the regularized one to the same equations. It is concluded that the exact non-homogeneous term can successfully be reconstructed.

Keywords
truncated singular value decomposition, regularized solution, Moore-Penrose pseudo-inverse matrix, exact non-homogeneous term

1 Introduction

We consider the following simultaneous linear equations:\(^4\):

\[
\begin{align*}
0.4x_1 + 1.2x_2 &= 5.2 \\
3.5x_1 + 10.501x_2 &= 45.504
\end{align*}
\]

(1)

and

\[
\begin{align*}
0.4x_1 + 1.2x_2 &= 5.2 \\
3.5x_1 + 10.501x_2 &= 45.501
\end{align*}
\]

(2)

The solution to Eq. (1) is \(x_1 = 1, x_2 = 4\), while the solution to Eq. (2) is \(x_1 = 10, x_2 = 1\). We can see that a small change of the right-hand side called the non-homogeneous term encompasses a large change of the solution.

The Cauchy problem of the Laplace equation\(^5\), known as an inverse problem of partial differential equations, is an ill-posed problem\(^5\). The resultant simultaneous linear equations after discretization of the Cauchy problem have the same issue as Eq. (1) (or Eq. (2)). Namely, a small noise contained in
the non-homogeneous term of the resultant equations affects sensitively on the accuracy of the solution. The non-homogeneous term consists of the observed data called the Cauchy data. For the Cauchy data without noises, an accurate numerical solution to the Cauchy problem can be obtained from the solution to the resultant simultaneous equations. On the other hand, for the noisy Cauchy data, a numerical solution to the Cauchy problem drastically jumbles due to a small noise contained in the non-homogeneous term. Then, applying a regularization method\(^1\) to the resultant simultaneous linear equations, we can obtain an accurate numerical solution to the Cauchy problem\(^3\).

Consequently, when we consider simultaneous linear equations like Eq. (1) (or Eq. (2)), we notice that the non-homogeneous term can be regarded as an exact one which does not contain noises if the solution obtained without applying the regularization method almost coincides with the solution obtained by applying the regularization method.

We therefore have no criterion for the correctness of the non-homogeneous term of arbitrarily given simultaneous linear equations. In this study, we will define an exact non-homogeneous term and propose how to reconstruct the exact one by using the truncated singular value decomposition as a regularization method.

### 2 Problem setting

For simplicity, we consider the following simultaneous linear equations:

\[
\begin{align*}
x_1 + ax_2 &= b, \\
x_1 + (a + \varepsilon)x_2 &= b + \delta,
\end{align*}
\]

where \(a, b\) are given real constants, \(\varepsilon > 0\) a given sufficiently small constant, and \(\delta > 0\) a given perturbed parameter. The simultaneous linear equations (3), (4) are written in the matrix form:

\[
Ax = b(\delta),
\]

where the coefficient matrix \(A\) and the both of the solution \(x\) and the non-homogeneous term \(b\) depending on \(\delta\) are denoted by

\[
A := \begin{pmatrix} 1 & a \\ 1 & a + \varepsilon \end{pmatrix}, \quad x = x(\delta) := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad b = b(\delta) := \begin{pmatrix} b \\ b + \delta \end{pmatrix}.
\]

Since \(A\) is not singular (\(|A| = \varepsilon \neq 0\)), the solution to Eqs. (3), (4) is obtained as

\[
x = x(\delta) = \begin{pmatrix} b \\ 0 \end{pmatrix} + \frac{\delta}{\varepsilon} \begin{pmatrix} -a \\ 1 \end{pmatrix}.
\]

If \(\delta\) changes to \(\delta'\), then the changes of the non-homogeneous term and the solution are given by

\[
\Delta b := b(\delta') - b(\delta) = \begin{pmatrix} 0 \\ \delta' - \delta \end{pmatrix},
\]

\[
\Delta x := x(\delta') - x(\delta) = \frac{\delta' - \delta}{\varepsilon} \begin{pmatrix} -a \\ 1 \end{pmatrix},
\]

respectively. A small change of the non-homogeneous term (\(|\Delta b| = |\delta' - \delta|\)) encompasses a large change of the solution if \(\delta' - \delta \geq \varepsilon\). Such simultaneous linear equations are called ill-conditioned. Therefore, we need to obtain the solution which does not drastically change for a small change of the non-homogeneous term.
3 Singular value decomposition (SVD)

Let $A^T$ denote the transpose of the matrix $A$. We consider the following eigenvalue problem of the matrix $A^T A$:

$$A^T A v = \lambda v, \quad v \neq 0,$$

where $\lambda$ denotes the eigenvalue of $A^T A$ and $v$ the eigenvector associated with $\lambda$. Eq. (6) is rewritten as

$$(A^T A - \lambda I)v = 0, \quad v \neq 0. \quad (7)$$

Since Eq. (7) has a non-zero solution $v$, we know that the determinant of $A^T A - \lambda I$ is zero. Hence, we obtain the characteristic equation of $A$:

$$|A^T A - \lambda I| = \lambda^2 - (a^2 + (a + \varepsilon)^2 + 2)\lambda + \varepsilon^2 = 0. \quad (8)$$

The eigenvalues $\lambda_1, \lambda_2$ ($\lambda_1 > \lambda_2$) are the solutions to Eq. (8), which satisfies

$$\lambda_1 + \lambda_2 = a^2 + (a + \varepsilon)^2 + 2 = 2(a^2 + 1) + O(\varepsilon), \quad (9)$$

$$\lambda_1\lambda_2 = \varepsilon^2, \quad (10)$$

where we assume $\varepsilon \to 0$ to derive the asymptotic expansion although $\varepsilon$ is a sufficiently small constant. The eigenvalue $\lambda_1$ is expressed in the form:

$$\lambda_1 = \frac{\{a^2 + (a + \varepsilon)^2 + 2\} + \sqrt{(a^2 + (a + \varepsilon)^2 + 2)^2 - 4\varepsilon^2}}{2}$$

$$= 2(a^2 + 1 + a\varepsilon) + O(\varepsilon^2) = 2(a^2 + 1) + O(\varepsilon)$$

(see Section A). Eq. (10) implies

$$\lambda_2 = \frac{\varepsilon^2}{\lambda_1} = \frac{\varepsilon^2}{2(a^2 + 1) + O(\varepsilon)} = \frac{\varepsilon^2}{2(a^2 + 1)} + O(\varepsilon^3).$$

The normalized eigenvectors associated with the eigenvalues $\lambda_1, \lambda_2$ are given as

$$v_1 := \frac{1}{\sqrt{a^2 + 1 + O(\varepsilon)}} \left( \begin{array}{c} 1 + O(\varepsilon) \\ a + O(\varepsilon) \end{array} \right) = \frac{1}{\sqrt{a^2 + 1}} \left( \begin{array}{c} 1 \\ a \end{array} \right) + O(\varepsilon),$$

$$v_2 := \frac{1}{\sqrt{a^2 + 1 + O(\varepsilon)}} \left( \begin{array}{c} a + O(\varepsilon) \\ -1 + O(\varepsilon) \end{array} \right) = \frac{1}{\sqrt{a^2 + 1}} \left( \begin{array}{c} a \\ -1 \end{array} \right) + O(\varepsilon),$$

respectively, where we abbreviate a vector or a matrix which consists of components $O(\varepsilon)$ as $O(\varepsilon)$ for simplicity. Since we see from $(A^T A)^T = A^T (A^T)^T = A^T A$ that $A^T A$ is symmetric, the eigenvectors $v_1, v_2$ of $A^T A$ are orthonormal.

Dividing Eq. (6) by $\sigma_i := \sqrt{\lambda_i}$, we have

$$A^T \left( \frac{1}{\sigma_i} Av_i \right) = \sigma_i v_i.$$

Let

$$u_i := \frac{1}{\sigma_i} Av_i,$$

- 3 -
which can be calculated as
\[ u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + O(\varepsilon), \]
\[ u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + O(\varepsilon). \]

Then, we have
\[ A^T u_i = \sigma_i v_i, \quad A v_i = \sigma_i u_i. \] (11)

Multiplying the first equation of Eq. (11) by the matrix \( A \) from the left side, we obtain
\[ A A^T u_i = A(\sigma_i v_i) = \sigma_i A v_i = \sigma_i(\sigma_i u_i) = \lambda_i u_i, \]
from which we know that \( u_i \) is the eigenvector associated with the eigenvalue \( \lambda_i \) of \( AA^T \). Since \( AA^T \) is symmetric, \( u_1, u_2 \) are orthonormal. We can confirm that \( u_1, u_2 \) are orthonormal also in this way:
\[
(u_i, u_j) = \left( \frac{1}{\sigma_i} A v_i, \frac{1}{\sigma_j} A v_j \right) = \frac{1}{\sigma_i \sigma_j} (A^T A v_i, v_j) = \frac{1}{\sigma_i \sigma_j} (\lambda_i v_i, v_j) \quad (\therefore A^T A v_i = \lambda_i v_i)
\]
\[
= \frac{\sigma_i}{\sigma_j} (v_i, v_j) \quad (\therefore \lambda_i = \sigma_i^2)
\]
\[ = \delta_{ij} \]
with the Kronecker delta \( \delta_{ij} \).

From the second equation of Eq. (11), we have
\[
(\sigma_1 u_1, \sigma_2 u_2) = (A v_1, A v_2)
\]
or
\[
(u_1, u_2) \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} = A(v_1, v_2).
\]

Since \((v_1, v_2)\) is an orthogonal matrix, multiplying the above equation by \((v_1, v_2)^T\) from the right side, we can obtain the singular value decomposition (SVD) of the matrix \( A \) as follows:
\[ A = U \Sigma V^T = \sum_{j=1}^{2} \sigma_j u_j v_j^T, \] (12)
where the orthogonal matrices \( U, V \) and the diagonal matrix \( \Sigma \) are given by
\[
U := (u_1, u_2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} + O(\varepsilon),
\]
\[
V := (v_1, v_2) = \frac{1}{\sqrt{a^2 + 1}} \begin{pmatrix} 1 & a \\ a & -1 \end{pmatrix} + O(\varepsilon),
\]
\[
\Sigma := \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} = \begin{pmatrix} \sqrt{2(a^2 + 1)}/2 & 0 \\ 0 & \varepsilon/\sqrt{2(a^2 + 1)} \end{pmatrix} + O(\varepsilon),
\]
and \( \sigma_j \) denotes the singular value of \( A \), and \( u_j, v_j \) the left singular vector and the right singular vector associated with \( \sigma_j \), respectively.
The condition number corresponding to 2-norm can be expressed as

\[
\text{cond}(A) = \frac{\sigma_1}{\sigma_2} = \left( \frac{\lambda_1}{\lambda_2} \right)^{1/2} = \left( \frac{\lambda_1}{\varepsilon^2/\lambda_1} \right)^{1/2} = \frac{\lambda_1}{\varepsilon} = 2(a^2 + 1 + a\varepsilon) + O(\varepsilon^2) \]

\[
= \frac{2(a^2 + 1)}{\varepsilon} + 2a + O(\varepsilon)
\]

with the maximal and the minimal singular values \( \sigma_1 \) and \( \sigma_2 \). We can see that \( \text{cond}(A) \) becomes large as the determinant \(|A| = \varepsilon > 0\) tends to zero.

4 Regularized solution by the truncated SVD

Based on the SVD, the inverse matrix of \( A \) can be expressed in the form:

\[
A^{-1} = (U\Sigma V^T)^{-1} = (V^T)^{-1}\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^T = \sum_{j=1}^{2} \frac{1}{\sigma_j} v_j u_j^T
\]

(13)

by using \( VV^T = I, UU^T = I \).

The solution \( x \) to \( Ax = b \) can be written in the form:

\[
x = A^{-1}b = \sum_{j=1}^{2} \frac{(u_j, b)}{\sigma_j} v_j
\]

and the solution \( x + \Delta x \) to \( A(x + \Delta x) = b + \Delta b \), that is, \( \Delta x \) to \( A\Delta x = \Delta b \) can be written in the form:

\[
\Delta x = A^{-1}\Delta b = \sum_{j=1}^{2} \frac{(u_j, \Delta b)}{\sigma_j} v_j,
\]

from which we can see that the second term \( (j = 2) \) of the right-hand side may be large since \( \sigma_2 \) is close to zero. That is the reason why the solution \( x \) retains the potential to change drastically for a small change of the non-homogeneous term \( b \).

Since the singular value \( \sigma_2 = \varepsilon/\sqrt{2(a^2 + 1)} \) is sufficiently close to zero in Eq. (12), the matrix \( A \) can be approximated as

\[
A_1 := \sigma_1 u_1 v_1^T = \begin{pmatrix} a \\ 1 \end{pmatrix} + O(\varepsilon),
\]

which is called the truncated SVD (TSVD). According to Eq. (13), the Moore-Penrose pseudo-inverse matrix of \( A_1 \) is defined by

\[
A_1^\dagger := \frac{1}{\sigma_1} v_1 u_1^T = \frac{1}{2(a^2 + 1)} \begin{pmatrix} 1 & 1 \\ a & a \end{pmatrix} + O(\varepsilon).
\]

Then, a regularized solution can be defined as

\[
x_1(\delta) := A_1^\dagger b(\delta) = \frac{(u_1, b(\delta))}{\sigma_1} v_1 = \frac{b}{a^2 + 1} \begin{pmatrix} 1 \\ a \end{pmatrix} + O(\varepsilon) + O(\delta),
\]

(14)

which does not have the term with sufficiently small \( \sigma_2 \) and hence does not drastically change for the perturbation of \( \delta \).
We now define an exact non-homogeneous term by $b(\delta)$ such that the first term of the right-hand side in the regularized solution (14) coincides with the conventional exact one (5). If the first term of the right-hand side in Eq. (14) equals Eq. (5), then we obtain

$$\delta = \delta_1 := \frac{ab\epsilon}{a^2 + 1},$$

which gives the exact non-homogeneous term as follows:

$$b(\delta_1) = \frac{b}{a^2 + 1} \left( \frac{a^2 + 1}{a^2 + 1 + a\epsilon} \right).$$

5 Least norm solution

The general solution to Eq. (3) is given by

$$x_g = \begin{pmatrix} b \\ 0 \end{pmatrix} + t \begin{pmatrix} -a \\ 1 \end{pmatrix}, \quad \forall t \in \mathbb{R}. \quad (16)$$

From

$$\|x_g\|^2 = (b - at)^2 + t^2 = (a^2 + 1) \left( t - \frac{ab}{a^2 + 1} \right)^2 + \frac{b^2}{a^2 + 1},$$

we know that

$$t = \frac{ab}{a^2 + 1} \quad (17)$$

implies the least norm solution

$$x_0 = \frac{b}{a^2 + 1} \begin{pmatrix} 1 \\ a \end{pmatrix}, \quad (18)$$

which coincides with the first term of the right-hand side in the regularized solution (14) except the sufficiently small term with respect to $\epsilon$ and $\delta$.

On the other hand, comparing the general solution (16) with the solution (5), we have

$$t = \frac{\delta}{\epsilon}. \quad (19)$$

From Eqs. (17) and (19), we can easily obtain (15).

In addition, the general solution to Eq. (4) is obtained as

$$x_g = \begin{pmatrix} b \\ 0 \end{pmatrix} + t \begin{pmatrix} -a \\ 1 \end{pmatrix} + \begin{pmatrix} \delta - t\epsilon \\ 0 \end{pmatrix}, \quad \forall t \in \mathbb{R},$$

which coincides with the solution (5) as $\delta - t\epsilon = 0$. Hence, we can confirm Eq. (19).

6 Solution norm against perturbation

The norm of the solution (5) is written as

$$\|x(\delta)\|^2 = \left( \frac{b - \delta}{\epsilon a} \right)^2 + \left( \frac{\delta}{\epsilon} \right)^2 = \frac{a^2 + 1}{\epsilon^2} \left( \frac{\delta - ab\epsilon}{a^2 + 1} \right)^2 + \frac{b^2}{a^2 + 1}. \quad (20)$$
Hence, if
\[ \delta = \delta_1 = \frac{ab\varepsilon}{a^2 + 1}, \]
then the solution norm \( \|x(\delta)\| \) depending on \( \delta \) is minimized. Consequently, the exact non-homogeneous term can be given by \( b(\delta_1) \) when we take \( \delta = \delta_1 \) which minimizes the solution norm \( \|x(\delta)\| \).

Putting \( \delta = \delta_1 \) in Eq. (5), we obtain the exact solution \( x \) to \( Ax = b \) with the exact \( b = b(\delta_1) \) as follows:
\[
x(\delta_1) = \begin{pmatrix} b \\ 0 \end{pmatrix} + \frac{ab}{a^2 + 1} \begin{pmatrix} -a \\ 1 \end{pmatrix} = \frac{b}{a^2 + 1} \begin{pmatrix} 1 \\ a \end{pmatrix}.
\]
(21)
Then, we can see that the regularized solution (14) can be regarded as an approximation to the exact solution (21) as follows:
\[
x_1(\delta_1) = x(\delta_1) + O(\varepsilon) + O(\delta_1).
\]

7 Numerical example

Let \( a = 5, b = 26 \) and \( \varepsilon = 10^{-3} \). We numerically obtain the conventional exact and the regularized solutions to the simultaneous linear equations (3), (4) for given \( \delta \).

<table>
<thead>
<tr>
<th>( \delta )</th>
<th>( x_1(\delta) )</th>
<th>( x(\delta) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 2 \cdot 10^{-3} )</td>
<td>(0.9998, 4.9997)(^T)</td>
<td>(16, 2)(^T)</td>
</tr>
<tr>
<td>( 5 \cdot 10^{-3} )</td>
<td>(0.9999, 5.0000)(^T)</td>
<td>(1, 5)(^T)</td>
</tr>
<tr>
<td>( 7 \cdot 10^{-3} )</td>
<td>(0.9999, 5.0002)(^T)</td>
<td>(−9, 7)(^T)</td>
</tr>
</tbody>
</table>

Table 1: Regularized solutions vs. conventional exact ones

Figure 1: Solution norm against \( \delta \)
Table 1 shows that the regularized solutions $x_1(\delta)$ for $\delta = 2 \cdot 10^{-3}, 5 \cdot 10^{-3}, 7 \cdot 10^{-3}$ are approximately equal to $(1, 5)^T$, while the conventional exact solutions $x(\delta)$ for $\delta = 2 \cdot 10^{-3}, 7 \cdot 10^{-3}$ are $(16, 2)^T, (-9, 7)^T$, respectively, which are much different from the regularized solutions. The conventional exact solution $x(\delta)$ for $\delta = 5 \cdot 10^{-3}$ is $(1, 5)^T$, which almost coincides with the regularized solution. Actually, we can see that if we take $\delta = \delta_1 = 5 \cdot 10^{-3}$ according to Eq. (15), the conventional exact solution almost coincides with the regularized one. We can therefore reconstruct the exact non-homogeneous term as $b(5 \cdot 10^{-3}) = (26, 26.005)^T$.

Figure 1 shows the solution norms $\|x(\delta)\|$ for $\delta = 2 \cdot 10^{-3}, 5 \cdot 10^{-3}, 7 \cdot 10^{-3}$ indicated by the small circles and the solid curve obtained by (20) against $\delta$ ($1 \cdot 10^{-3} \leq \delta \leq 9 \cdot 10^{-3}$). We can see from the figure that $\|x(\delta)\|$ is minimized at $\delta = \delta_1 = 5 \cdot 10^{-3}$ which gives the exact non-homogeneous term.

8 Conclusions

In this study, we have considered the ill-conditioned simultaneous linear equations in two unknowns. Since the non-homogeneous term of simultaneous linear equations appearing in real engineering problems consists of observed noisy data, the exact non-homogeneous term without noises is unknown. The exact non-homogeneous term based on the TSVD is defined and reconstructed in the study. This approach will be extended to reconstruct exact data from observed noisy data in ill-conditioned simultaneous linear equations in $n$ unknowns.

A Appendix

For any $p \in \mathbb{R}$, the binomial series

$$ (x + h)^p = \sum_{k=0}^{\infty} \binom{p}{k} h^k x^{p-k} $$

holds for any $h$ such that $|h| < |x|$ with the generalized binomial coefficients

$$ \binom{p}{k} := \frac{p(p-1) \cdots (p-k+1)}{k!}. $$

Then, we have

$$ (x + h)^p = x^p + O(h) \quad (h \to 0). \quad (22) $$

Using Eq. (22), we can derive

$$ \sqrt{\{a^2 + (a + \varepsilon)^2 + 2\}^2 - 4\varepsilon^2} = \sqrt{\{2(a^2 + 1 + a\varepsilon) + \varepsilon^2\}^2 - 4\varepsilon^2} \quad (\therefore x = 2(a^2 + 1 + a\varepsilon), \ h = \varepsilon^2, \ p = 2 \ in \ Eq. \ (22)) $$

$$ = \sqrt{\{2(a^2 + 1 + a\varepsilon)\}^2 + O(\varepsilon^2)} \quad (\therefore x = \{2(a^2 + 1 + a\varepsilon)\}^2, \ h = O(\varepsilon^2), \ p = 1/2 \ in \ Eq. \ (22)) $$

$$ = 2(a^2 + 1 + a\varepsilon) + O(\varepsilon^2) \quad (\therefore a^2 + 1 + a\varepsilon = (a + \varepsilon/2)^2 + 1 - \varepsilon^2/4 > 0) \quad (23) $$

$$ = 2(a^2 + 1) + O(\varepsilon) \quad (24) $$
as $\varepsilon \to 0$, where the asymptotic expansion (23) is more accurate than the expansion (24). Therefore, we obtain

$$
\lambda_1 = \frac{\{a^2 + (a + \varepsilon)^2 + 2\} + \{2(a^2 + 1 + a\varepsilon) + O(\varepsilon^2)\}}{2}
= 2(a^2 + 1 + a\varepsilon) + O(\varepsilon^2) = 2(a^2 + 1) + O(\varepsilon).
$$

References


