New choices for Tikhonov regularization matrix using fractional derivative approach

Somaieh Mohammady[1] and Mohammad Reza Eslahchi[1]

[1]Tarbiat Modares University

April 16, 2024

Abstract

A linear discrete ill-posed problem has a perturbed right-hand side vector and an ill-conditioned coefficient matrix. The solution to such a problem is very sensitive to perturbation. Replacement of the coefficient matrix by a nearby one that has less condition number is one of the well-known approaches for decreasing the sensitivity of the problem to perturbation. In this paper, we suggest some new regularization matrix to the Tikhonov regularization. These new ones are based on fractional derivatives such as Grunwald-Letnikov and Caputo and can cause to have more exact solutions.

New choices for Tikhonov regularization matrix using fractional derivative approach

Somaieh Mohammady[*] M. R. Eslahchi[†]

Department of Applied Mathematics, Faculty of Mathematical Sciences, Tarbiat Modares University, P.O.Box 14115-134, Tehran, Iran.

November 19, 2021

Abstract

A linear discrete ill-posed problem has a perturbed right-hand side vector and an ill-conditioned coefficient matrix. The solution to such a problem is very sensitive to perturbation. Replacement of the coefficient matrix by a nearby one that has less condition number is one of the well-known approaches for decreasing the sensitivity of the problem to perturbation.

In this paper, we suggest some new regularization matrix to the Tikhonov regularization. These new ones are based on fractional derivatives such as Grunwald-Letnikov and Caputo and can cause to have more exact solutions.

Keywords : Ill-posed problem, Regularization matrix, Tikhonov regularization.

MSC : $65F22$, $47A52$, $65R30$.

1 Introduction

Researchers have found that a great number of problems can be classified as ill-posed problems. These problems arise in quantum mechanics, astronomy, optimal control theory, image deblurring [15], ultrasound testing, etc [21]. A linear discrete ill-posed problem is the discretization of an ill-posed problem such as the Fredholm integral equation of the first kind. This problem is a linear system in which the coefficient matrix is ill-conditioned and the right-hand side vector is contaminated by error [16]. The noise-free system is not available. The ill-conditioning of the coefficient matrix leads to propagation of the noise in the solution. Therefore, a straightforward solution to this system is not meaningful. So the system must be regularized.

Regularization is the replacement of the available system by a nearby one that is less sensitive to perturbation. The goal of regularization is to obtain an approximated solution close to the exact solution of the unavailable noise-free system. There are some methods for regularization. They can be classified as direct methods and iterative methods. Some of the direct methods are Truncated singular value decomposition (TSVD) [18, 33] and Tikhonov regularization [32] and some of the iterative methods are Krylov subspace-based methods such as CGLS, LSQR

[*]Somaieh Mohammady (E-mail address: s.mohammady@modares.ac.ir)

[†]Corresponding author. M. R. Eslahchi (E-mail address: eslahchi@modares.ac.ir).

and GMRES [2, 10, 19, 26, 27]. Tikhonov regularization (TR) is one of the most popular methods of regularization.
In this method, a regularized term is added to the available system. This term has a matrix called the regularization matrix. The regularized solution of the TR method depends definitively on the regularization matrix. A good choice of this matrix can help to approximate the solution in a more stable way. In this paper, we will focus on the TR matrix and introduce some new matrices based on fractional derivatives.
For the readers' convenience, the main contributions of this paper are highlighted as follows:

- Some new regularization matrices are introduced based on Grunwald-Letnikov and Caputo fractional derivative operators.
- The new matrices are dependent on a parameter α. A proper choice of this parameter gives the user a more appropriate solution.
- Some properties of the new matrices are shown in the next sections and an important one is proved in detail (see Theorem 3).
- Numerical examples are implemented to demonstrate the efficiency and accuracy of the new matrices and the comparisons are given by the other regularization ones based on other derivative operators.

The organization of this paper is as follows. Section 2 is devoted to study the Tikhonov regularization (TR) methods and we review some researches on TR matrices. Section 3 introduces some new TR matrices based on fractional derivatives. The advantages of these new matrices are shown by solving some numerical examples in Section 4.

2 Tikhonov regularization and regularization matrix

Consider the following least-square problem:

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|, \quad A \in \mathbb{R}^{m\times n}, \quad m \geq n. \tag{1}$$

In this paper, $\|.\|$ denote the Euclidean vector norm. The matrix A is ill-conditioned in which the singular values gradually decay to zero with no particular gap. The matrix A may be singular and the right-hand side b is as follows:

$$b = \overline{b} + e, \tag{2}$$

where e is an unknown noise vector. The unknown error-free system

$$Ax = \overline{b}, \tag{3}$$

is consistent and its solution is denoted by $\overline{x}$. Since (3) is not available, we want to approximate $\overline{x}$ by solving the available linear system (1). The ill-conditioning of A causes the propagation of the noise of b in the solution and straightforward methods for solving this system does not yield a meaningful solution. So researchers replace the available system (1) by a nearby one which is less sensitive to the noise in data. This approach is called regularization. One of the most popular methods of regularization is Tikhonov regularization (TR). In this method, a regularized term is added to (1). The general form of TR is as follows:

$$\min_{x\in\mathbb{R}^n} \|Ax-b\|^2+\lambda^2\|Lx\|^2, \tag{4}$$

in which L is called regularization matrix and λ is the regularization parameter. The regularization parameter determines how well x_λ approximates $\overline{x}$ and how x_λ is sensitive to the error e in the available data b. Note that there are different methods for the determination of regularization parameter. Some of them are based on an estimation of $\|e\|$, such as the discrepancy principle. In this method, λ is determined such that

$$\|Ax_\lambda-b\|=\eta\delta, \tag{5}$$

where δ is an estimation of $\|e\|$, $\eta\geq 1$ is an user-specified constant independent of δ and close to unity [12, 16]. When there is no estimation of noise, one can choose generalized cross validation, L-curve criterion, etc. For more details, we refer the reader to [12, 16, 28, 30]. Before getting into regularization matrices, let's review some basic principles.

Definition 1. *The null space of the matrix $A_{m\times n}$ is denoted $N(A)$ and is defined:*

$$N(A)=\{x\in\mathbb{R}^n|Ax=0\}.$$

Theorem 1. *If the assumption*

$$N(A)\cap N(L)=\{0\},$$

is satisfied, then the solution of (4) *by normal equation, i.e.,*

$$x_\lambda=(A^TA+\lambda^2L^TL)^{-1}A^Tb, \tag{6}$$

is unique [16].

Note that in the standard form of TR, L is equal to I, i.e.,

$$\min_{x\in\mathbb{R}^n}\{\|Ax-b\|^2+\lambda^2\|x\|^2\}, \tag{7}$$

Theorem 2. *(Singular value decomposition) Suppose A be an $m\times n$ matrix. Then there exists a factorization, called singular value decomposition (SVD) of A of the form*

$$A=U\Sigma V^T, \tag{8}$$

where $\Sigma=diag[\sigma_1,\sigma_2,\cdots,\sigma_n]\in R^{m\times n}$, $\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_r>\sigma_{r+1}=\cdots=\sigma_n=0$, $rank(A)=r$, $U=[u_1,u_2,\cdots,u_m]\in R^{m\times m}$, $U^TU=I$, $V=[v_1,v_2,\cdots,v_n]\in R^{n\times n}$ and $V^TV=I$.

Proof. See [14]. $\square$

The solution of the standard form of TR (7) by normal equation is

$$x_\lambda = (A^T A + \lambda^2 I)^{-1} A^T b, \tag{9}$$

and by using the SVD of matrix A in (9), the regularized solution of (7) could be written as

$$x_{reg} = \sum_{i=1}^{r} \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \frac{u_i^T b}{\sigma_i} v_i. \tag{10}$$

The general form of TR (4) can be transformed to the standard TR [16]. For doing this, let the A-weighted pseudoinverse of matrix L be defined by

$$L_A^\dagger := (I - (A(I - L^\dagger L))^\dagger A) L^\dagger,$$

where $L^\dagger$ denotes the Moore-Penrose pseudoinverse of L. Now consider

$$x_0 = (A(I - L^\dagger L))^\dagger b,$$

and

$$\hat{A} := A L_A^\dagger, \quad \hat{b} := b - A x_0, \quad \hat{x} := Lx.$$

So the minimization problem (4) can be transformed to the following standard form [10]:

$$\min_{\hat{x} \in \mathbb{R}^n} \left\| \hat{A}\hat{x} - \hat{b} \right\|^2 + \lambda^2 \|\hat{x}\|^2. \tag{11}$$

The solution x_λ in (6) can be achieved by the solution $\hat{x}_\lambda$ of (11) as follows

$$x_\lambda = L_A^\dagger \hat{x}_\lambda + x_0. \tag{12}$$

Note that the regularization scheme does not affect on the component x_0. Indeed, x_0 belongs to the null space of L (i.e., $N(L)$) [16].

Remark 1. *If the matrix L is square and non-singular, then by considering $\hat{A} = AL^{-1}, \hat{b} = b$ and $\hat{x} = Lx$ in* (11)*, the transformation of* (4) *to standard TR is very simple [16].*

According to (6), the regularization matrix has a direct effect on the regularized solution. So a good choice of regularization matrix can further improve the regularized solution. For this aim, researchers are interested to introduce more relevant matrices. Here we look at some of the proposed ones. To the best of the author's knowledge, regularization matrices can be classified into two types. One type is non-derivative-based matrices and another is derivative-based ones. In the first type, some of the matrices are about multiplying a diagonal matrix by an orthogonal one. In [13], the authors propose the regularization matrix

$$L = DV^T, \tag{13}$$

in which D is a diagonal matrix consisting of singular values of A and regularization parameter λ and V^T obtained by the SVD of A in (8). They consider D_λ^2 as follows

$$D_\lambda^2 = diag[\max\{\lambda^2 - \sigma_1^2, 0\}, \max\{\lambda^2 - \sigma_2^2, 0\}, \cdots, \max\{\lambda^2 - \sigma_n^2, 0\}].$$

Some other diagonal matrices for (13) are suggested in [5, 25, 34]. In [24], an extension form of diagonal matrix D in (13) by adding some parameters to the Tikhonov regularization method is considered. The worthy point of this matrix is that it is an extended form of all the presented methods in [13, 25, 34] with appropriate choices of parameters. The added parameters are obtained by solving a linear fractional programming problem based on some conditions.
Now, the second type is introduced. It can be stated that one of the most desirable TR matrices is based on the discrete approximation to derivative operators.
To show this type of matrices, let us consider the interval $[a, b]$ and define the mesh points

$$x_j = a + jh, \quad j = 0, 1, ..., n, \tag{14}$$

where $h = \frac{b-a}{n}$.
The approximation of m-th order derivative operators for $m \in \mathbb{N}$ in the mesh points (14) is as follows:

$$f^{(m)}(x_j) = \frac{d^m f(x_j)}{dx^m} = \lim_{h \to 0} \frac{1}{h^m} \sum_{r=0}^{m} (-1)^r \binom{m}{r} f(x_{j+r}), \quad j = 0, 1, ..., n-m, \tag{15}$$

and the banded matrix L_m

$$L_m = \begin{bmatrix} l_0 & l_1 & l_2 & \cdots & l_m & & & 0 \\ & l_0 & l_1 & l_2 & \cdots & l_m & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & & & l_0 & l_1 & l_2 & \cdots & l_m \end{bmatrix}_{(n-m)\times n}, \quad l_i = (-1)^i \binom{m}{i}, \quad i = 0, 1, ..., m, \tag{16}$$

is the scaled approximations to the m-th order derivative. Common choices of regularization matrices are L_1 and L_2 as follows:

$$L_1 = \begin{bmatrix} 1 & -1 & & & 0 \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & -1 \end{bmatrix}_{(n-1)\times n}, \tag{17}$$

and

$$L_2 = \begin{bmatrix} 1 & -2 & 1 & & & 0 \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ 0 & & & 1 & -2 & 1 \end{bmatrix}_{(n-2)\times n}, \tag{18}$$

which are obtained by (16) for $m = 1$ and $m = 2$, respectively.
The null space of L_1 and L_2 are

$$N(L_1) = span\{[1, 1, \cdots, 1]^T\},$$

and

$$N(L_2) = span\{[1, 1, \cdots, 1]^T, [1, 2, 3, \cdots, n]^T\}.$$

respectively [16].
Calvetti et al. construct invertible smoothing preconditioners from L_1 and L_2 which are well suited for use with the iterative methods [3]. The authors in [7] discuss the design of square regularization matrices based on L_1 and L_2 that can be used in iterative methods based on the Arnoldi process for large-scale Tikhonov regularization problems. With a focus on the boundary conditions, some square regularization matrices from finite difference equations are introduced in [8]. Reichel and Ye [29] present two approaches for constructing some square regularization matrices: zero-padding of the rectangular matrices and extending the rectangular matrix to a square circulant. The authors in [20] consider the distance of a matrix to the closest matrix with a user-specified null space by using the Frobenius norm and describe an approach to construct some square regularization matrices. Dykes et al. in [9] extends the approach of [20] to problems in higher space dimensions.
In the next section, we suggest a new approach to constructing TR matrices by fractional derivative operators.

3 The new regularization matrices based on fractional derivatives

Due to the high usage of L_1 and L_2 as TR matrices, the authors of this work are focused on fractional derivatives. Firstly, consider the following definition:

Definition 2. *For a function f defined on an interval $J = [a, b]$, the Grunwald-Letnikov fractional derivative of order $\alpha > 0$ is defined by [22]:*

$$ {}_G D^{\alpha}_{x,b} f(x) = \lim_{h\to 0} \quad \frac{1}{h^{\alpha}} \sum_{r=0}^{[\frac{b-x}{h}]} (-1)^r \binom{\alpha}{r} f(x + rh). \tag{19}$$

So (19) can lead to the following approximation in the mesh points (14)

$$ {}_G D^{\alpha}_{x_j,b} f(x_j) = \frac{1}{h^{\alpha}} \sum_{r=0}^{n-j} (-1)^r \binom{\alpha}{r} f(x_{j+r}). \tag{20}$$

To use the Grunwald-Letnikov fractional derivative operator by (20), this paper suggests the following case as TR matrix which include the scaled coefficients of (20) in the mesh points:

$$ {}_G L_{\alpha} := \begin{bmatrix} (-1)^0\binom{\alpha}{0} & (-1)^1\binom{\alpha}{1} & (-1)^2\binom{\alpha}{2} & \cdots & (-1)^{n-1}\binom{\alpha}{n-1} \\ 0 & (-1)^0\binom{\alpha}{0} & (-1)^1\binom{\alpha}{1} & \cdots & (-1)^{n-2}\binom{\alpha}{n-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ & & & (-1)^0\binom{\alpha}{0} & (-1)^1\binom{\alpha}{1} \\ 0 & \cdots & & 0 & (-1)^0\binom{\alpha}{0} \end{bmatrix}_{n\times n}. \tag{21}$$

It is remarkable that this square matrix is non-singular and this causes the simplicity of transformation to the standard form of TR (see Remark 1) and the null space of a non-singular

matrix only contains the zero vector.
Note that for $\alpha = 1$, ${}_GL_1$ is as follows (If $\alpha \in \mathbb{N}$ and $r > \alpha$, then $\binom{\alpha}{r} = 0$):

$$
{}_GL_1 = \begin{bmatrix} 1 & -1 & & & 0 \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ 0 & & & & 1 \end{bmatrix}_{n\times n}, \tag{22}
$$

which is equal to the suggested square matrix in [7] for approximation of the first derivative and for $\alpha = 2$, ${}_GL_2$ is as follows:

$$
{}_GL_2 = \begin{bmatrix} 1 & -2 & 1 & & & 0 \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \\ 0 & & & & & 1 \end{bmatrix}_{n\times n}, \tag{23}
$$

which could be use as a square regularization matrix for approximation of the second derivative. Generally, ${}_GL_1$ and ${}_GL_2$ are similar to (17) and (18) for $n \times n$ case, respectively.
Now we consider another fractional derivative formula:

Definition 3. *For a function f defined on an interval $J = [a, b]$, the Caputo fractional derivative of order $\alpha > 0$ is defined by [22]:*

$$
{}_CD^{\alpha}_{x,b}f(x) = \frac{(-1)^N}{\Gamma(N-\alpha)} \int_x^b \frac{f^{(N)}(t)}{(t-x)^{\alpha-N+1}} dt, \tag{24}
$$

where $N = [\alpha] + 1$.

Consider the discretization of (24) by using finite difference method for $0 < \alpha < 1$ ($N = 1$) in the mesh points (14) which is proposed in [23]:

$$
{}_CD^{\alpha}_{x_j,b}f(x_j) = \frac{1}{\Gamma(2-\alpha)h^{\alpha}} \sum_{r=j}^{n-1} (f_{r+1} - f_r) g_{j,r}, \tag{25}
$$

where $g_{j,r} = (r-j)^{1-\alpha} - (r-j+1)^{1-\alpha}$.
In order to use the Caputo fractional derivative operator for $0 < \alpha < 1$, It is suggested the following regularization matrix:

$$
{}_CL^1_{\alpha} = \begin{bmatrix} -g_{1,1} & h_{1,2} & h_{1,3} & \cdots & h_{1,n-1} & g_{1,n-1} \\ 0 & -g_{2,2} & h_{2,3} & \cdots & h_{2,n-1} & g_{2,n-1} \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ & & & -g_{n-2,n-2} & h_{n-2,n-1} & g_{n-2,n-1} \\ 0 & & \cdots & 0 & -g_{n-1,n-1} & g_{n-1,n-1} \end{bmatrix} \in \mathbb{R}^{(n-1)\times n}, \tag{26}
$$

in which

$$h_{i,j} = g_{i,j-1} - g_{i,j}. \tag{27}$$

The discretization of (24) by using finite difference method for $1 < \alpha < 2$ ($N = 2$) in the mesh points (14) is as follows [23]:

$$_C D^{\alpha}_{x_j,b} f(x_j) = \frac{1}{\Gamma(3-\alpha)h^{\alpha}} \sum_{r=j}^{n-1} (f_{r+1} - 2f_r + f_{r-1}) c_{j,r}, \tag{28}$$

where

$$c_{j,r} = (r-j+1)^{2-\alpha} - (r-j)^{2-\alpha}. \tag{29}$$

The following regularization matrix for utilizing the Caputo fractional derivative with $1 < \alpha < 2$ is proposed:

$$_C L^2_{\alpha} = \begin{bmatrix} c_{2,2} & d_{2,2} & q_{2,3} & q_{2,4} & \cdots & q_{2,n-2} & l_{2,n-2} & c_{2,n-1} \\ 0 & c_{3,3} & d_{3,3} & q_{3,4} & & & l_{3,n-2} & c_{3,n-1} \\ \vdots & & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ & & & c_{n-3,n-3} & d_{n-3,n-3} & q_{n-3,n-2} & l_{n-3,n-2} & c_{n-3,n-1} \\ & & & & c_{n-2,n-2} & d_{n-2,n-2} & l_{n-2,n-2} & c_{n-2,n-1} \\ 0 & & & \cdots & 0 & c_{n-1,n-1} & -2c_{n-1,n-1} & c_{n-1,n-1} \end{bmatrix} \in \mathbb{R}^{(n-2)\times n}, \tag{30}$$

where

$$d_{i,j} = -2c_{i,j} + c_{i,j+1}, \quad q_{i,j} = c_{i,j-1} - 2c_{i,j} + c_{i,j+1}, \quad l_{i,j} = c_{i,j} - 2c_{i,j+1}. \tag{31}$$

Since Theorem 1 plays an important role on the regularized solution, the null spaces of $N(_C L^1_{\alpha})$ and $N(_C L^2_{\alpha})$ are studied in the following theorem.

Theorem 3. *The null spaces of* $_C L^1_{\alpha}$ *and* $_C L^2_{\alpha}$ *are*

$$N(_C L^1_{\alpha}) = N(L_1) = span\{[1, 1, \cdots, 1]^T\},$$

and

$$N(_C L^2_{\alpha}) = N(L_2) = span\{[1, 1, \cdots, 1]^T, [1, 2, 3, \cdots, n]^T\},$$

respectively.

Proof. Suppose $x = [x_1, x_2, ..., x_n]^T$ and consider $_C L^1_{\alpha} x = 0$. Now we solve the last equation (i.e., $(n-1)$th one) of $_C L^1_{\alpha} x = 0$:

$$-g_{(n-1)(n-1)} x_{n-1} + g_{(n-1)(n-1)} x_n = 0,$$

which leads to $x_n = x_{n-1}$.
By choosing the $(n-2)$th equation of $_C L^1_{\alpha} x = 0$, i.e.,

$$-g_{(n-2)(n-2)} x_{n-2} + h_{(n-2)(n-1)} x_{n-1} + g_{(n-2)(n-1)} x_n = 0,$$

and by substitution (27) in $h_{(n-2)(n-1)}$, then it is concluded that $x_{n-2} = x_n$.
Now consider the $(n-i)$th equation $(1 \leq i \leq n-1)$:

$$\begin{aligned}&- g_{(n-i)(n-i)}x_{n-i} + h_{(n-i)(n-i-1)}x_{n-i-1} + h_{(n-i)(n-i-2)}x_{n-i-2} + ...\\&+ h_{(n-i)(n-1)}x_{n-1} + g_{(n-2)(n-1)}x_n = 0,\end{aligned}$$

and by (27) we conclude that $x_{n-i} = x_n$.
Hence the solution of ${}_CL^1_\alpha x = 0$ is $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} x_n$.
and the null space of ${}_CL^1_\alpha$ is spanned by $[1, 1, \ldots, 1]^T$.
Now to find the null space of ${}_CL^2_\alpha$, consider The $(n-2)$th equation of ${}_CL^2_\alpha x = 0$:

$$c_{(n-1)(n-1)}x_{n-2} - 2c_{(n-1)(n-1)}x_{n-1} + c_{(n-1)(n-1)}x_n = 0,$$

and it can be obtained directly that

$$x_{n-2} = 2x_{n-1} - x_n. \tag{32}$$

Next, the $(n-3)$th equation of ${}_CL^2_\alpha x = 0$ is

$$c_{(n-2)(n-2)}x_{n-3} + d_{(n-2)(n-2)}x_{n-2} + l_{(n-2)(n-2)}x_{n-1} + c_{(n-2)(n-1)}x_n = 0,$$

and by substitution (31) in above equation, it can be seen that

$$x_{n-3} = 3x_{n-1} - 2x_n. \tag{33}$$

We repeat the substitution (31) in the $(n-4)$th equation of ${}_CL^2_\alpha x = 0$, i.e.,

$$c_{(n-3)(n-3)}x_{n-4} + d_{(n-3)(n-3)}x_{n-3} + q_{(n-3)(n-2)}x_{n-2} + l_{(n-3)(n-2)}x_{n-1} + c_{(n-3)(n-1)}x_n = 0,$$

and it leads to $x_{n-4} = 4x_{n-1} - 3x_n$.
Generally, the $(n-i)$th equation for $i = 5, \ldots, (n-1)$ of ${}_CL^2_\alpha x = 0$ is:

$$\begin{aligned}&c_{(n-i+1)(n-i+1)}x_{n-i} + d_{(n-i+1)(n-i+1)}x_{n-i+1} + q_{(n-i+1)(n-i+2)}x_{n-i+2} +\\&q_{(n-i+1)(n-i+3)}x_{n-i+3} + ... + q_{(n-i+1)(n+2)}x_{n-2} + l_{(n-i+1)(n-i+2)}x_{n-1} +\\&c_{(n-i+1)(n-1)}x_n = 0,\end{aligned}$$

and substitution (31) in the above equation leads to

$$x_{n-i} = ix_{n-1} - (i-1)x_n.$$

Therefore, the solution of ${}_CL^2_\alpha x = 0$ is

$$\begin{bmatrix} x_1 \\ \vdots \\ x_{n-3} \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} (n-1)x_{n-1} - (n-2)x_n \\ \vdots \\ 3x_{n-1} - 2x_n \\ 2x_{n-1} - x_n \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} n-1 \\ \vdots \\ 3 \\ 2 \\ 1 \\ 0 \end{bmatrix} x_{n-1} + \begin{bmatrix} -(n-2) \\ \vdots \\ -2 \\ -1 \\ 0 \\ 1 \end{bmatrix} x_{n-1}.$$

This shows that the null space of ${}_CL_\alpha^2$ is spanned by $w_1 = [n-1, n-2, \ldots, 1, 0]^T$ and $w_2 = [1, 1, \cdots, 1]^T$.
Now we want to show $span\{w_1, w_2\} = span\{v_1, v_2\}$ where $v_1 = [1, 2, \cdots, n]^T$ and $v_2 = [1, 1, \cdots, 1]^T$. For this purpose, it is sufficient to show that v_1 and v_2 can be written as linear combination of w_1 and w_2 and vice versa.
It can be easily concluded that

$$\begin{cases} w_1 = nv_2 - v_1, \\ w_2 = v_1 - (n-1)v_2, \end{cases} \qquad \begin{cases} v_1 = (n-1)w_1 + nw_2, \\ v_2 = w_1 + w_2. \end{cases}$$

So the null space of ${}_CL_\alpha^2$ is spanned by the vectors $[1, 2, \cdots, n]^T$ and $[1, 1, \cdots, 1]^T$ $\square$

From $N({}_CL_\alpha^1) = N(L_1)$ and $N({}_CL_\alpha^2) = N(L_2)$, it is concluded that the unregularized components of x_0 of x_λ (12) for L_1 and ${}_CL_\alpha^1$ are the same. The same is true for L_2 and ${}_CL_\alpha^2$. To have more insight into the TR matrices, Figure 1 shows the graphs of $||{}_CL_\alpha^1 x||^2 = 1, \quad 0 < \alpha < 1$ and $||{}_GL_{0.5}^1 x||^2 = 1$ in blue and green lines, respectively (for $n = 2$). It is interesting that the graphs of $||{}_CL_\alpha^1 x||^2 = 1, \quad 0 < \alpha < 1$ and $||L_1 x||^2 = 1$ are the same.

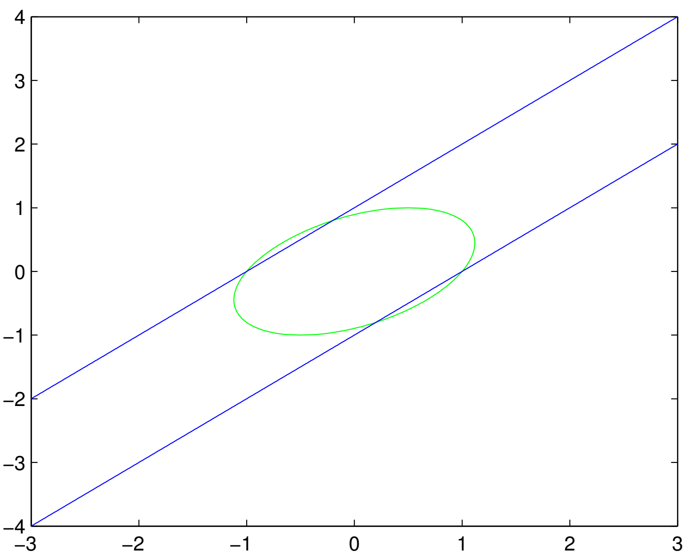

Figure 1: $||{}_GL_{0.5}x||^2 = 1$ by green line and $||{}_CL_\alpha^1 x||^2 = 1, \quad 0 < \alpha < 1$ by blue line.

Remark 2. *The ${}_GL_\alpha$ matrix is the extension of all the previous derivative-based TR matrices, i.e., L_1 and L_2.*

Remark 3. *The ${}_CL_\alpha^1$ and ${}_CL_\alpha^2$ have the same properties with L_1 and L_2, respectively, such as the order of matrix and the null space. With regard to Caputo fractional derivative formula (24), it can be seen that the first and second derivative are used in $0 < \alpha < 1$ and $1 < \alpha < 2$, respectively. Therefore, it seems convenient to compare the numerical results of ${}_CL_\alpha^1$ and ${}_GL_\alpha$ (for $0 < \alpha \leq 1$) with L_1 and the results of ${}_CL_\alpha^2$ and ${}_GL_\alpha$ (for $1 < \alpha \leq 2$) with L_2.*

4 Numerical examples

We illustrate the performance of the regularization matrices (21), (26) and (30) and compare them to the first and second derivative-based regularization matrices (17)-(18). The computations are carried out in Matlab software. Consider the error-free linear system

$$A\bar{x} = \bar{b},$$

in which $A \in \mathbb{R}^{n\times n}, \bar{x} \in \mathbb{R}^n, \bar{b} \in \mathbb{R}^n$. Note that $\bar{b}$ and $\bar{x}$ are the error-free right-hand side and the exact solution respectively. A white Gaussian noise vector $e \in \mathbb{R}^m$ with zero mean is added to $\bar{b}$ according to (2) to yield the right-hand side b. The noise vector e is scaled to obtain a specified noise level defined by

$$\varepsilon = \frac{\|e\|}{\|\bar{b}\|}.$$

The relative error is computed by $\frac{\|x^{app}-\bar{x}\|}{\|\bar{x}\|}$ in which x^{app} shows the approximated solution of TR method (4) with different regularization matrices. For simplicity, we call Grunwald-Letnikov and Caputo derivative-based regularization matrices as GR and CR, respectively. All the ill-posed problems in examples are taken from the MATLAB package Regularization Tools [17].

Example 4.1. The first test problem is an integral equation which is studied in [31] by **Shaw** as follows:

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} k(s,t)f(t)dt = g(s), \quad -\frac{\pi}{2} \leq s \leq \frac{\pi}{2}, \tag{34}$$

where the kernel k is given by

$$k(s,t) = (cos(s) + cos(t))^2 (sin(u)/u)^2, \quad u = \pi\left(sin\left(s\right) + sin\left(t\right)\right).$$

The exact solution is

$$f(t) = 2exp(-6(t-\frac{4}{5})^2) + exp(-2(t+\frac{1}{2})^2).$$

The discretization of $k(s,t)$ and $f(t)$ by simple quadrature lead to produce $A \in \mathbb{R}^{n\times n}$ and $\bar{x} \in \mathbb{R}^n$, respectively. Then the discrete right-hand side $\bar{b}$ is produced as $\bar{b} := A\bar{x}$. Table 1 displays the results over 20 runs for noise level 1% and $n = 100$.

The results in Table 1 show that GR and CR (fractional derivative-based regularization) matrices have better performance than L_1 and L_2. To have more insight into the accuracy of different regularization matrices, Figure 2 shows the approximated solutions and $\bar{x}$.

Example 4.2. The Fredholm integral equation of the first kind of problem **Deriv2** discussed in [6] is

$$\int_0^1 k(s,t)f(t)dt = g(s), \quad 0 \leq s \leq 1. \tag{35}$$

TR matrix	α	Relative error	TR matrix	α	Relative error
${}_GL_\alpha$	0.2	$1.83 \cdot 10^{-1}$	${}_GL_\alpha$	1.2	$1.73 \cdot 10^{-1}$
	0.4	$2.01 \cdot 10^{-1}$		1.4	$1.57 \cdot 10^{-1}$
	0.6	$2.02 \cdot 10^{-1}$		1.6	$1.42 \cdot 10^{-1}$
	0.8	$1.97 \cdot 10^{-1}$		1.8	$1.31 \cdot 10^{-1}$
	1	$1.86 \cdot 10^{-1}$		2	$1.31 \cdot 10^{-1}$
${}_CL^1_\alpha$	0.2	$2.53 \cdot 10^{-1}$	${}_CL^2_\alpha$	1.2	$2.27 \cdot 10^{-1}$
	0.4	$2.47 \cdot 10^{-1}$		1.4	$2.23 \cdot 10^{-1}$
	0.6	$2.42 \cdot 10^{-1}$		1.6	$2.19 \cdot 10^{-1}$
	0.8	$2.36 \cdot 10^{-1}$		1.8	$2.14 \cdot 10^{-1}$
L_1	–	$5.43 \cdot 10^{-1}$	L_2	–	$6.53 \cdot 10^{-1}$

Table 1: Average relative error of **Shaw** test problem with noise level 1% and $n = 100$ (Example 4.1).

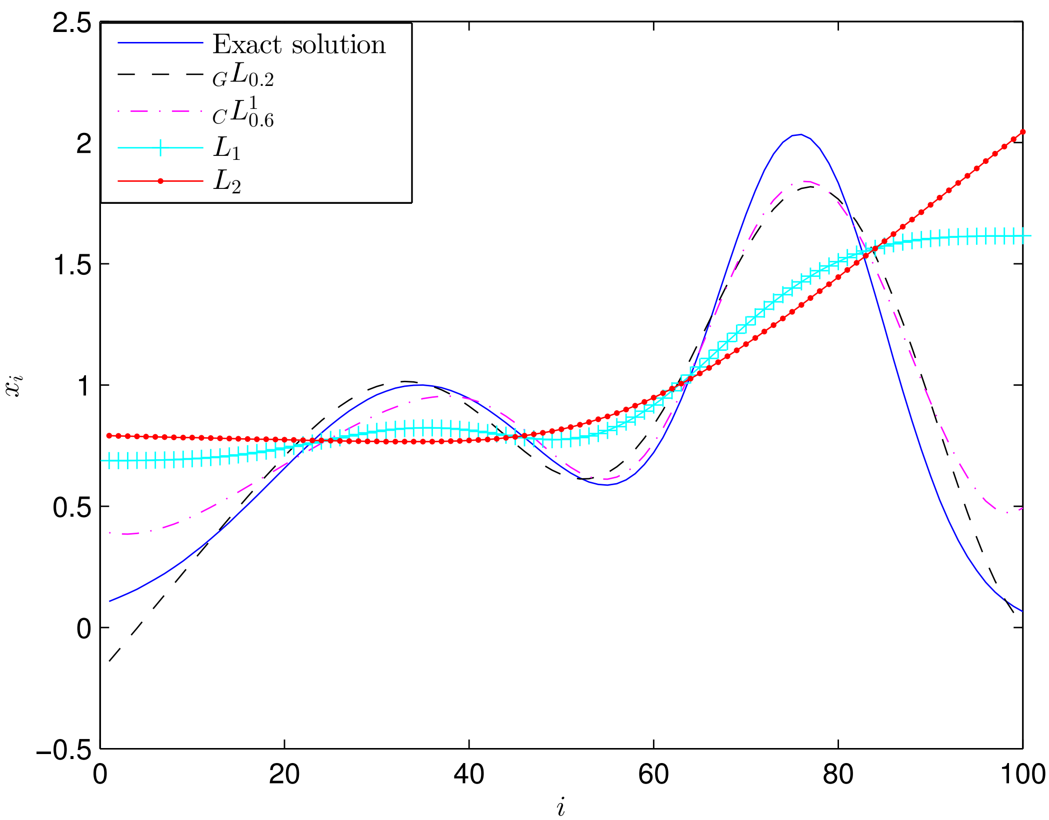

Figure 2: Approximated solutions with ${}_GL_{0.2}$, ${}_CL^1_{0.6}$, L_1, L_2 and Exact solution for **Shaw** test problem with noise level 1% and $n = 100$ (Example 4.1).

The kernel is

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

where

$$f(t)=exp(t),$$

and the right-hand side g is given by

$$g(s)=exp(s)+(1-e)s-1.$$

This problem is discretized by Galerkin method with orthonormal box functions which leads to produce a linear system with the coefficient matrix $A\in\mathbb{R}^{n\times n}$ and the right-hand side vector $\bar{b}\in\mathbb{R}^n$.

The average relative errors of the approximated solutions over 20 runs for noise level 50% with $n=100$ are shown in Table 2. Again, GR and CR matrices have less relative errors and give more accurate results than the results of L_1 and especially L_2. Figure 3 shows the approximated solutions and $\bar{x}$.

TR matrix	α	Relative error	TR matrix	α	Relative error
${}_G L_\alpha$	0.2	$3.98\cdot10^{-1}$	${}_G L_\alpha$	1.2	$5.04\cdot10^{-1}$
	0.4	$3.93\cdot10^{-1}$		1.4	$5.30\cdot10^{-1}$
	0.6	$4.17\cdot10^{-1}$		1.6	$5.54\cdot10^{-1}$
	0.8	$4.48\cdot10^{-1}$		1.8	$5.81\cdot10^{-1}$
	1	$4.77\cdot10^{-1}$		2	$6.12\cdot10^{-1}$
${}_C L^1_\alpha$	0.2	$2.33\cdot10^{-1}$	${}_C L^2_\alpha$	1.2	$2.54\cdot10^{-1}$
	0.4	$2.27\cdot10^{-1}$		1.4	$2.54\cdot10^{-1}$
	0.6	$2.19\cdot10^{-1}$		1.6	$2.54\cdot10^{-1}$
	0.8	$2.08\cdot10^{-1}$		1.8	$2.54\cdot10^{-1}$
L_1	–	$4.71\cdot10^{0}$	L_2	–	$1.02\cdot10^{+2}$

Table 2: Average relative error of **Deriv2** test problem with noise level 50% and $n=100$ (Example 4.2).

Example 4.3. Consider the following first kind of Fredholm integral equation

$$\int_0^\pi k(s,t)f(t)dt=g(s),\quad 0\leq s\leq\frac{\pi}{2}, \tag{36}$$

with kernel k and right-hand side g given by

$$k\left(s,t\right)\ =\ exp\left(scos\left(t\right)\right),\quad g\left(s\right)\ =\ 2sinh\left(s\right)/s.$$

The exact solution is

$$f\left(t\right)=sin\left(t\right).$$

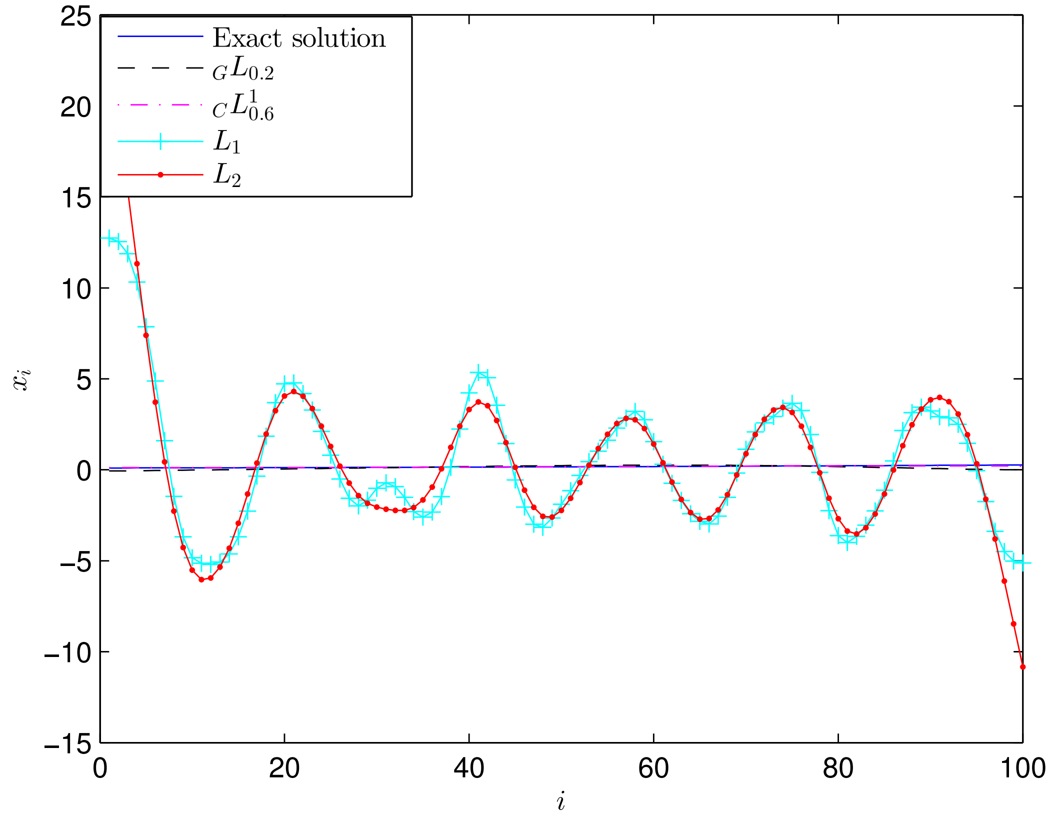

Figure 3: Approximated solutions with $_GL_{0.2}$, $_CL^1_{0.6}$, L_1, L_2 and Exact solution for **Deriv2** test problem with noise level 50% and $n = 100$ (Example 4.1).

This integral equation is discussed by **Baart** [1].
It is discretized by the Galerkin method to produce a linear system $A\bar{x} = \bar{b}$ with the coefficient matrix $A \in \mathbb{R}^{n\times n}$ and the right-hand side vector $\bar{b} \in \mathbb{R}^n$.

Table 3 shows the average relative errors over 20 runs with noise level 5% and $n = 100$.

Example 4.4. Consider **Heat** [4, 11] the inverse heat equation which is a Volterra integral equation of the first kind on $[0, 1]$ as integration interval with kernel $k(s, t) = K(s - t)$, where

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

The exact solution is taken to be

$$x(t) = \begin{cases} \frac{300}{4}t^2 & 0 \le t \le \frac{1}{10}, \\ \frac{3}{4} + (20t - 2)(3 - 20t) & \frac{1}{10} < t \le \frac{3}{20}, \\ \frac{3}{4}e^{2(3-20t)} & \frac{3}{20} < t \le \frac{1}{2}, \\ 0 & \frac{1}{2} < t \le 1. \end{cases}$$

TR matrix	α	Relative error	TR matrix	α	Relative error
${}_G L_\alpha$	0.2	$1.95 \cdot 10^{-1}$	${}_G L_\alpha$	1.2	$9.83 \cdot 10^{-2}$
	0.4	$1.65 \cdot 10^{-1}$		1.4	$7.74 \cdot 10^{-2}$
	0.6	$1.46 \cdot 10^{-1}$		1.6	$6.40 \cdot 10^{-2}$
	0.8	$1.32 \cdot 10^{-1}$		1.8	$7.68 \cdot 10^{-1}$
	1	$1.17 \cdot 10^{-1}$		2	$1.08 \cdot 10^{-1}$
${}_C L_\alpha^1$	0.2	$2.77 \cdot 10^{-1}$	${}_C L_\alpha^2$	1.2	$2.66 \cdot 10^{-1}$
	0.4	$2.74 \cdot 10^{-1}$		1.4	$2.61 \cdot 10^{-1}$
	0.6	$2.70 \cdot 10^{-1}$		1.6	$2.53 \cdot 10^{-1}$
	0.8	$2.64 \cdot 10^{-1}$		1.8	$2.43 \cdot 10^{-1}$
L_1	–	$4.38 \cdot 10^{-1}$	L_2	–	$6.86 \cdot 10^{0}$

Table 3: Average relative error of **Baart** test problem with noise level 5% and $n = 100$ (Example 4.3).

Discretization of this problem is done by means of simple quadrature (midpoint rule) and a linear system with the coefficient matrix $A \in \mathbb{R}^{n\times n}$ and exact solution $\bar{x} \in \mathbb{R}^n$ is produced. The right-hand side vector $\bar{b} \in \mathbb{R}^n$ is obtained by $\bar{b} := A\bar{x}$.

The average relative errors of the approximated solutions over 20 runs for noise level 0.01% and $n = 100$ are shown in Table 4.

TR matrix	α	Average relative error	TR matrix	α	Average relative error
${}_G L_\alpha$	0.2	$1.83 \cdot 10^{-2}$	${}_G L_\alpha$	1.2	$1.84 \cdot 10^{-2}$
	0.4	$1.81 \cdot 10^{-2}$		1.4	$1.86 \cdot 10^{-2}$
	0.6	$1.82 \cdot 10^{-2}$		1.6	$1.87 \cdot 10^{-2}$
	0.8	$1.83 \cdot 10^{-2}$		1.8	$1.88 \cdot 10^{-2}$
	1	$1.84 \cdot 10^{-2}$		2	$1.89 \cdot 10^{-2}$
${}_C L_\alpha^1$	0.2	$5.77 \cdot 10^{-2}$	${}_C L_\alpha^2$	1.2	$1.89 \cdot 10^{-2}$
	0.4	$4.37 \cdot 10^{-2}$		1.4	$1.95 \cdot 10^{-2}$
	0.6	$2.56 \cdot 10^{-2}$		1.6	$1.97 \cdot 10^{-2}$
	0.8	$1.87 \cdot 10^{-2}$		1.8	$1.99 \cdot 10^{-2}$
L_1	–	$5.30 \cdot 10^{-1}$	L_2	–	$6.82 \cdot 10^{-1}$

Table 4: Average relative error of **Heat** test problem with noise level 0.01% and $n = 100$ (Example 4.4).

Example 4.5. Here an image restoration example is provided by **Blur** from the regularization tool package [17]. The image is contaminated by 0.1% noise level and the parameters $\sigma = 0.7$ and *band* $= 3$ is considered for the blurring function. The performance of different TR matrices is shown by images and relative errors.

Table 5 shows that ${}_G L_\alpha$ could not produce good results. However, ${}_C L_\alpha$ yields the least relative errors and the restored images by this matrix are very clear than those obtained by L_1 and L_2.

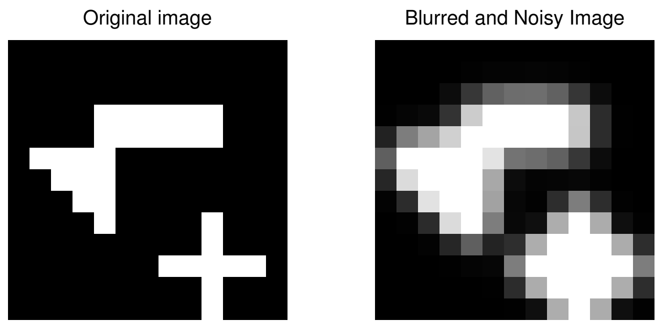

Figure 4: True image and blurred and noisy one (Example 4.5).

TR matrix	α	Relative error	TR matrix	α	Relative error
${}_GL_\alpha$	0.2	$4.40 \cdot 10^{-2}$	${}_GL_\alpha$	1.2	$8.99 \cdot 10^{-1}$
	0.4	$4.33 \cdot 10^{-2}$		1.4	$9.00 \cdot 10^{-1}$
	0.6	$8.94 \cdot 10^{-1}$		1.6	$9.02 \cdot 10^{-1}$
	0.8	$9.04 \cdot 10^{-1}$		1.8	$9.05 \cdot 10^{-1}$
	1	$9.10 \cdot 10^{-1}$		2	$9.08 \cdot 10^{-1}$
${}_CL^1_\alpha$	0.2	$1.25 \cdot 10^{-2}$	${}_CL^2_\alpha$	1.2	$1.43 \cdot 10^{-2}$
	0.4	$1.38 \cdot 10^{-2}$		1.4	$1.44 \cdot 10^{-2}$
	0.6	$1.39 \cdot 10^{-2}$		1.6	$1.44 \cdot 10^{-2}$
	0.8	$1.40 \cdot 10^{-2}$		1.8	$1.45 \cdot 10^{-2}$
L_1	–	$4.19 \cdot 10^{-2}$	L_2	–	$8.79 \cdot 10^{-1}$

Table 5: Relative error of **Blur** test problem with noise level 0.1% (Example 4.5).

5 Conclusion

Using derivative operators is the common way as Tikhonov regularization matrix. In this paper, the authors extend this approach and introduce some new Tikhonov regularization matrices based on fractional derivative. One of the advantages of this extension is the existence of parameter α in these new matrices. The good choice of this parameter can help to decrease the relative error and so to improve the regularized solution. Another interesting point for these research is obtaining the null spaces of matrices, explicitly. The results show that the new matrices have superiority in reducing the relative errors over the classic derivative-based regularization matrices (first and second derivatives). The future work of the authors is the study of other fractional derivative operators which could be utilizable as TR matrices.

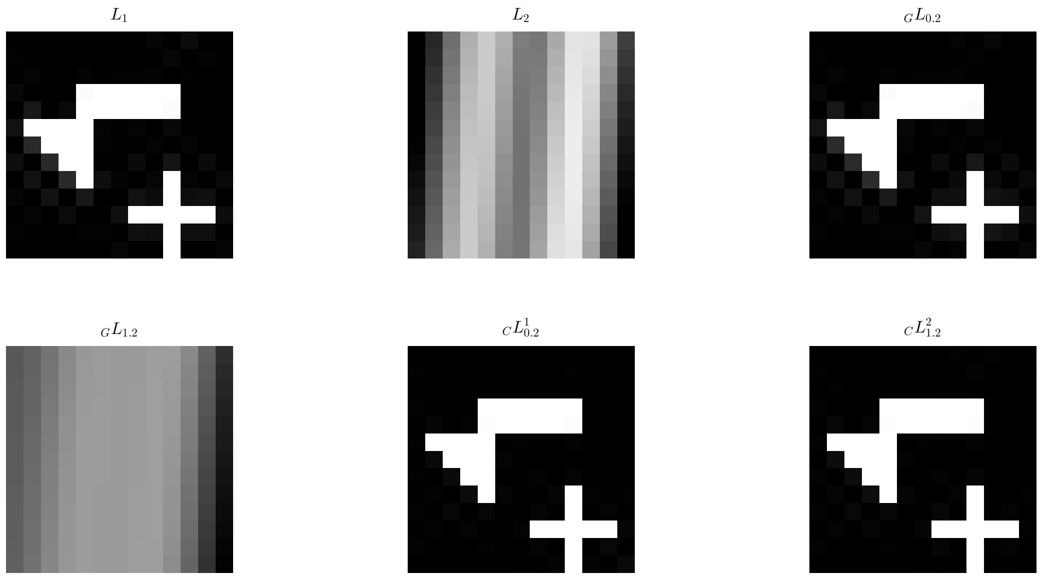

Figure 5: Restored images with different TR matrices (Example 4.5).

References

[1] M. L. Baart, The use of auto-correlation for pseudo-rank determination in noisy ill-conditioned least squres problems, IMA J. Numer. Anal., 2 (1982), pp. 241-247.

[2] A. Bjorck, Numerical Methods for Least Squares Problems, SIAM, Philadelphia, 1996.

[3] D. Calvetti, L. Reichel, and A. Shuibi, Invertible smoothing preconditioners for linear discrete ill-posed problems, Appl. Numer. Math., 54 (2005), pp. 135–149.

[4] A. S. Carasso, Determining surface temperatures from interior observations, SIAM J. Appl. Math., 42 (1982), pp. 558-574.

[5] J. Cui, G. Peng, Q. Lu and Z. Huang, A special modified Tikhonov regularization matrix for discrete ill-posed problems, Appl. Math. Comput., 377 (2020), 10.1016/j.amc.2020.125165.

[6] L. M. Delves, J. L. Mohamed, "Computational Methods for Integral Equations", Cambridge University Press, 1985, p. 310.

[7] M. Donatelli, A. Neuman, L. Reichel, Square regularization matrices for large linear discrete ill-posed problems, Numer. Linear Algebra Appl., 19 (2012), pp. 896–913.

[8] M. Donatelli, L. Reichel, Square smoothing regularization matrices with accurate boundary conditions, J. Comput. Appl. Math., 272 (2014), pp. 334–349.

[9] L. Dykes, G. Huang, S. Noschese, L. Reichel, Regularization matrices for discrete ill-posed problems in several space dimensions, Numer Linear Algebra Appl., (2018), 10.1002/nla.2163.

[10] L. Eldén, A weighted pseudoinverse, generalized singular values, and constrained least squares problems, BIT 22 (1982) 487–501.

[11] L. Eldén, The numerical solution of a non-characteristic Cauchy problem for a parabolic equation; in P. Deuflhart, E. Hairer (Eds.), Numerical Treatment of Inverse Problems in Differential and Integral Equations, BirkhÄauser, Boston, 1983.

[12] H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Kluwer, Dordecht, 1996.

[13] M. Fuhry, L. Reichel, A new Tikhonov regularization method, J. Numer. Algorithms., 59 (2012), pp. 433-445.

[14] G.H. Golub, C.F. Van Loan, Matrix Computations, third ed., the Johns Hopkins University Press, Baltimore, MD, 1996.

[15] P. C. Hansen, J. G. Nagy, and D. P. O'Leary, Deblurring Images, Matrices, Spectra, and Filtering, Fundamentals of Algorithms 3, SIAM, 2006.

[16] P. C. Hansen, Rank-Deficient and Discrete Ill-Posed Problems. Numerical Aspects of Linear Inversion, SIAM., Philadelphia, 1997.

[17] P. C. Hansen, Regularization tools version 4.0 for Matlab 7.3, *Numer. Algorithms.,* 46 (2007), pp. 189-194.

[18] R. J. Hanson, A numerical method for solving Fredholm integral equations of the first kind using singular values, SIAM J. Numer. Anal., 8 (1971), pp. 616-622.

[19] M. R. Hestenes and E. Stiefel, Methods of conjugate gradients for solving linear systems, J. Res. Nat. Bur. Standards, 49 (1952), pp. 409-436.

[20] G. Huang, S. Noschese, L. Reichel, Regularization matrices determined by matrix nearness problems, Linear Algebra Appl., 502 (2015), 502, pp. 41-57.

[21] S. I. Kabanikhin, Definitions and examples of inverse and ill-posed problems, J. Inv. Ill-Posed Problems, 16 (2008), 317-357.

[22] A. Kilbas, M.H. Srivastava, J.J. Trujillo, Theory and Application of Fractional Differential Equations, in: North Holland Mathematics Studies, vol. 204(2006).

[23] C. Li, F. Zeng, Finite difference methods for fractional differential equations, International Journal of Bifurcation and Chaos, 22 (2012), 10.1142/S0218127412300145.

[24] S. Mohammady , M.R. Eslahchi, Extension of Tikhonov regularization method using linear fractional programming, J. Comput. Appl. Math., 371 (2020), 10.1016/j.cam.2019.112677.

[25] S. Noschese, L. Reichel, Some matrix nearness problems suggested by Tikhonov regularization, Linear Algebra Appl., 502 (2015), pp. 366-386.

[26] C. C. Paige and M. A. Saunders, LSQR: An algorithm for sparse linear equations and sparse least squares, ACM Thans. Math. Software, 8 (1982), pp.43-71.

[27] C. C. Paige and M. A. Saunders, Algorithm 583. LSQR: Sparse linear equations and least squares problems, ACM Thans. Math. Software, 8 (1982), pp. 195-209.

[28] L. Reichel, H. Sadok, A new L-curve for ill-posed problems, J. Comput. Appl. Math., 219 (2008), pp. 493-508.

[29] L. Reichel, Q. Ye, Simple square smoothing regularization operators, Electron. Trans. Numer. Anal., 33 (2009), pp. 63–83.

[30] M. Rezghi, S. M. Hosseini, A new variant of L-curve for Tikhonov regularization, J. Comput. Appl. Math., 231 (2009), pp. 914-924.

[31] C. B. Shaw, Jr., Improvements of the resolution of an instrument by numerical solution of an integral equation, J. Math. Anal. Appl., 37 (1972), pp. 83-112.

[32] A. N. Tikhonov, Solution of incorrectly formulated problems and the regularization method, Soviet Math. Dokl., 4 (1963), pp. 1035-1038; English translation of Dokl. Akad. Nauk. SSSR, 151 (1963), pp. 501-504.

[33] J. M. Varah, On the numerical solution of ill-conditioned linear systems with applications to ill-posed problems, SIAM J. Numer. Anal., 10 (1973), pp. 257-267.

[34] X. J. Yang, L. Wang, A modified Tikhonov regularization method, J. Comput. Appl. Math., 288 (2015), pp. 180-192.