Research Article ISSN Online: 2783-4859 Computational Mathematics and Computer Modeling with Applications

Received 12 August 2024
Accepted 15 September 2024

DOI: 10.48308/CMCMA.3.1.21
AMS Subject Classification: 65F10; 65F22

Regularization properties of range restricted LSQR method for solving large-scale linear discrete ill-posed problems

Hui Zhang[a] and Hua Dai[b]

The LSQR iterative method is one of the most popular methods for solving large-scale linear discrete ill-posed problem $Ax = b$ with an error-contaminated right-hand side. In this paper, we consider the regularization properties of range restricted LSQR (RRLSQR) method. The iteration number k always acts as the regularization parameter because of the semi-convergence. In order to verify whether or not the RRLSQR method finds a 2-norm filtering best regularization solution for severely, moderately and mildly ill-posed problems, we present the $sin\Theta$ theorems for the 2-norm distances between the k dimensional left and right Krylov subspaces generated by Lanczos bidiagonalization and the k dimensional dominant left and right singular subspaces of A, and estimate the distances for the three kinds problems assuming that the singular values are simple, and develop a regularized RRLSQR method for solving linear discrete ill-posed problems. Numerical experiments confirm our theoretical results and show the efficiency of the proposed method.

Keywords: Linear discrete ill-posed problem; semi-convergence; range restricted LSQR method; regularization property.

1. Introduction

We are concerned with the large-scale linear discrete ill-posed problem

$$\min_{x\in R^n} \|Ax - b\| \quad \text{or} \quad Ax = b, \quad A \in \mathbf{R}^{m\times n}, b \in \mathbf{R}^m, \tag{1}$$

where $\|\cdot\|$ denotes the Euclidean vector norm or the spectral matrix norm, the matrix A is of ill-determined rank with its singular values decaying to zero with increasing index without a significant gap, the right-hand side $b = b_{true} + e$ represents the available error-contaminated data, b_{true} denotes the unavailable error-free vector associated with b and is assumed to be in the range of A, the available error-contaminated right-hand side b might not be, and e is the unknown noise, caused by measurement, modeling or discretization errors with $\|e\| < \|b_{true}\|$. We assume that a bound $\delta > 0$ of the norm of the noise e is known, i.e., $\|e\| \leq \delta$. For convenience, we also assume that the matrix A is square, i.e., $m = n$. In fact, this restriction can be removed since the least squares solution of (1) is equivalent to the solution of its normal equation $A^TAx = A^Tb$ whose coefficient matrix is square. Problem (1) arises from the discretization of inverse problems [29, 38], linear ill-posed problems [20, 25, 29] and the first kind of Fredholm integral equation [38, 39]. Many other applications, including pattern classification [11], inverse black body radiation [14], image restoration [41], computerized tomography [4, 43], dimensionality reduction [55] and so on, require also the solutions of problem (1). Due to the presence of the noise e in b and the severe ill-conditioning of the matrix A, the naive solution $x_{naive} = A^\dagger b$ to problem (1) does not furnish a useful approximation of the true solution $x_{true} = A^\dagger b_{true}$, where $A^\dagger$ denotes the Moore-Penrose generalized inverse of the matrix A. Therefore, regularization methods must be used to extract as good an approximation to x_{true} as possible. One of the most popular regularization methods is Tikhonov regularization [53], which replaces problem (1) by the following minimization problem

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

[a] Department of Basic Courses, Jiangsu Police Institute, Nanjing 210031, P.R. China.
[b] School of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P.R. China.
* Correspondence to: H. Dai. Email: hdai@nuaa.edu.cn

where $\lambda > 0$ is referred to as the regularization parameter.

Let

$$A = U\Sigma V^T, \tag{3}$$

be the singular value decomposition (SVD) of the matrix A, where $U = [u_1, u_2, \ldots, u_n] \in \mathbf{R}^{n\times n}$ and $V = [v_1, v_2, \ldots, v_n] \in \mathbf{R}^{n\times n}$ are orthogonal matrices, whose columns are the left and right singular vectors of A, respectively, $\Sigma = diag(\sigma_1, \sigma_2, \ldots, \sigma_n) \in \mathbf{R}^{n\times n}$ with the singular values $\sigma_1 > \sigma_2 > \ldots > \sigma_n > 0$ assumed to be simple, and the superscript T denotes the transpose of a vector or matrix. The behavior of problem (1) depends on the decay rate of the singular values σ_i of A. Problem (1) is severely ill-posed [29] if $\sigma_j = \mathcal{O}(\rho^{-j})(j = 1, 2, \ldots, n)$ with $\rho > 1$; problem (1) is mildly or moderately ill-posed [33, 34] if $\sigma_j = \mathcal{O}(j^{-\alpha})(i = 1, 2, \ldots, n)$ with $\frac{1}{2} < \alpha \le 1$ or $\alpha > 1$.

The naive and true solutions of problem (1) may be expressed respectively as

$$x_{naive} = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i = \sum_{i=1}^{n} \frac{u_i^T b_{true}}{\sigma_i} v_i + \sum_{i=1}^{n} \frac{u_i^T e}{\sigma_i} v_i,$$

and

$$x_{true} = \sum_{i=1}^{n} \frac{u_i^T b_{true}}{\sigma_i} v_i,$$

where $\|x_{true}\| = \|A^\dagger b_{true}\| \le C$ with some constant C, which means that b_{true} satisfies the discrete Picard condition [17, 25], i.e., $|u_i^T b_{true}|$ decays faster than σ_i. The common model is $|u_i^T b_{true}| = \sigma_i^{1+\beta}(i = 1, 2, \ldots, n)$ with $\beta > 0$ [25, 26]. The regularization solution x_λ of the minimization problem (2) is

$$x_\lambda = \sum_{i=1}^{n} f_i \frac{u_i^T b}{\sigma_i} v_i,$$

where $f_i = \frac{\sigma_i^2}{\sigma_i^2+\lambda^2}$ is a filter, and the optimal regularization parameter λ_{opt} should be determined such that $\|x_{\lambda_{opt}} - x_{true}\| = \min\limits_{\lambda>0} \|x_\lambda - x_{true}\|$. However, it is difficult to find the optimal regularization parameter λ_{opt} since b_{true} is not available. Some practical methods for choosing regularization parameters have been developed by using the discrepancy principle [42], the L-curve criterion [23, 40], and the generalized cross validation (GCV) [19].

Another effective and reliable regularization method for solving a small or medium sized ill-posed problem (1) is the truncated singular value decomposition (TSVD) method [22, 24, 25]. Let $U_k = [u_1, u_2, \ldots, u_k], V_k = [v_1, v_2, \ldots, v_k] \in \mathbf{R}^{n\times k}$ and $\Sigma_k = (\sigma_1, \sigma_2, \ldots, \sigma_k)$, then $A_k = U_k\Sigma_k V_k^T$ is the best rank k approximation to A under the spectral norm and $\|A - A_k\| = \sigma_{k+1}$ [3]. The TSVD method computes the TSVD regularization solution of problem (1) as follows

$$x_k^{tsvd} = A_k^\dagger b = \sum_{i=1}^{k} \frac{u_i^T b}{\sigma_i} v_i,$$

where the index k, determined by GCV [19], is a regularization parameter that determines how many dominant singular values and corresponding left and right singular vectors of A are used to compute a regularization solution.

If the noise e is the Gaussian white noise, then its covariance matrix is $\eta^2 I$, the expected value $\mathcal{E}(\|e\|^2) = n\eta^2$, $\mathcal{E}(|u_i^T e|) = \eta$ $(i = 1, 2, \ldots, n)$, and $\|e\| \approx \sqrt{n}\eta$ and $|u_i^T e| \approx \eta$ $(i = 1, 2, \ldots, n)$; see, e.g., [25, 26]. The positive integer number k_0 that satisfies

$$|u_{k_0}^T b| \approx |u_{k_0}^T b_{true}| > |u_{k_0}^T e| \approx \eta, \quad |u_{k_0+1}^T b| \approx |u_{k_0+1}^T e| \approx \eta.$$

is called the transition point [25, 26]. It follows from [26, p. 71, 86-8, 95] that k_0 is an optimal regularization parameter, i.e., k_0 satisfies $\|x_{k_0}^{tsvd} - x_{true}\| = \min\limits_{1\le k\le n} \|x_k^{tsvd} - x_{true}\|$, and $x_{k_0}^{tsvd}$ is the best TSVD regularized solution of problem (1). In order to assess the regularization ability of an Euclidean norm filtering regularization method, we can take $x_{k_0}^{tsvd}$ as the standard reference.

Besides the Tikhonov regularization method [53] and the TSVD method [22, 24, 25], which are either computationally unfeasible or extremely time-consuming for large-scale problems, iterative regularization has received considerable attention. Krylov iterative solvers are a major class of methods for solving problem (1). When A is symmetric positive definite, Gilyazov [18] and Hanke [20] developed the conjugate gradient (CG) method for solving the ill-posed problem (1), respectively. Plato [49] analyzed the regularizing properties of CG. When A is symmetric and indefinite, based on the Lanczos tridiagonalization, Paige and Saunders [47] proposed the MINRES method for solving indefinite systems of linear equations. Hanke [20], Kilmer and Stewart [37], Jensen and Hansen [32], Huang and Jia [30] analyzed regularizing effects of MINRES and showed its semi-convergence: the iterates first converge to x_{true}, then the noise e starts to deteriorate the iteration so that they start to diverge from x_{true}, and converge to x_{naive} instead. Therefore, the iteration number plays the role of the regularization parameter in iterative regularization methods. When A is nonsymmetric, Björck [2] presented the CGLS algorithm, which implicitly applies CG to the normal equations $A^T Ax = A^T b$. Hanke [20] studied the regularizing properties of CGLS. Based on the Lanczos bidiagonalization, Paige and Saunders [48] proposed the LSQR algorithm, which is mathematically equivalent to CGLS. The regularizing effects of the LSQR algorithm are analyzed in [31]. The GMRES method [50] is a popular iterative method for solving large linear systems of equations. Calvetti et al. [8] studied regularizing properties of the GMRES method. For iterative regularization, selecting a good stopping iteration is a crucial task. In order to overcome semi-convergence behavior and to avoid selecting a regularization parameter a priori, some hybrid methods based on LSQR, GMRES and Tikhonov regularization have been proposed; see, e.g., [2, 3, 8, 7, 9, 10, 12, 13, 16, 21, 27, 35, 36, 46].

Since the available error-contaminated right-hand side b might not be in the range of A, which causes the linear system $Ax = b$ to be inconsistent. Substituting Ab for b in Krylov subspace methods ensures that the projected system is consistent. It is natural to apply iterative methods with iterates in the range of A to the approximate solution of the ill-posed problem (1). Hanke [20] proposed a variant MR-II of the MINRES method for the solution of linear discrete ill-posed problems with a symmetric matrix A, which uses the starting vector Ab and restricts the resulting Krylov subspace to the range of A. The regularizing effects of the range restricted minimal residual Krylov subspace method are analyzed in [15, 20, 30]. Based on the Arnoldi process, Neuman et al. [44, 45] developed the range restricted GMRES methods for the solution of large nonsymmetric linear discrete ill-posed problems, which was originally designed for solving singular and inconsistent linear systems [5]. It has been observed in [6, 28] that the range restricted GMRES methods usually provides better regularized solutions than the GMRES method. Bellalij et al. [1] analyzed the properties of range restricted GMRES methods.

In this paper, motivated by the work in [44, 45], we develop the range restricted LSQR (RRLSQR) method for solving large-scale linear ill-posed problems (1). In order to analyze regularizing effects of the RRLSQR method, we assume that b_{true} and the transition point k_0 satisfy the following conditions

$$|v_i^T b_{true}| = \sigma_i^{1+\beta},\ \beta > 0,\ \ i = 1, 2, \dots, n, \tag{4}$$

and

$$|v_{k_0}^T b| \approx |v_{k_0}^T b_{true}| > |v_{k_0}^T e| \approx \eta, \quad |v_{k_0+1}^T b| \approx |v_{k_0+1}^T e| \approx \eta. \tag{5}$$

Our numerical experiments confirm that the inequality (5) is also true. We investigate the regularizing properties of the RRLSQR method, and establish some relationship between the underlying k dimensional left and right Krylov subspaces and the dominant left and right singular subspaces, which shows how the underlying k dimensional left and right Krylov subspaces approximate the dominant left and right singular subspaces, respectively. We develop a hybrid RRLSQR method that combines the RRLSQR method with a regularization method applied to the lower dimensional projected problems for solving large-scale linear ill-posed problems (1).

The paper is organized as follows. In Section 2, we present the RRLSQR method and analyze its regularizing effects. In Section 3, we establish the $sin\Theta$ theorems for the 2-norm distances between the underlying k dimensional left and right Krylov subspaces and the k dimensional dominant left and right singular subspaces of A, and derive accurate estimates on them for severely, moderately and mildly ill-posed problems, respectively. In Section 4, we develop a regularized RRLSQR method for solving large-scale linear ill-posed problems (1). In Section 5, we report some numerical experiments on test matrices to confirm our theoretical results and show the efficiency of the proposed method. Finally, some concluding remarks are given in Section 6.

Throughout this paper, we denote by $K_k(A, b) = span\{b, Ab, \dots, A^{k-1}b\}$ the k dimensional Krylov subspace generated by the matrix A and the vector b, and by I and the bold letter $\mathbf{0}$ the identity matrix and the zero matrix, respectively. Let $\mathcal{U}_k = span(U_k)$ and $\mathcal{V}_k = span(V_k)$ be the k dimensional dominant left and right singular subspaces of A, respectively. $\Theta(\mathcal{X}, \mathcal{Y})$ denotes the canonical angle between the two subspaces $\mathcal{X}$ and $\mathcal{Y}$ of the same dimension [52]. For the matrix $B = (b_{ij})$, define $|B| = (|b_{ij}|)$, and for $|C| = (|c_{ij}|)$, $|B| \le |C|$ means $|b_{ij}| \le |c_{ij}|$.

2. The RRLSQR method and its regularizing effects

The k-step Lanczos bidiagonalization process, which starts with an initial unit vector p_1 and computes two orthonormal bases $\{q_1, q_2, \dots, q_k\}$ and $\{p_1, p_2, \dots, p_{k+1}\}$ of the Krylov subspaces $K_k(A^TA, A^Tp_1)$ and $K_{k+1}(AA^T, p_1)$, respectively, may be described as follows.

Algorithm 1. The k-step Lanczos bidiagonalization process.

Input: $A \in R^{n\times n}$, an initial unit vector $p_1 \in R^n$, a positive integer k and set $\beta_1 q_0 = \mathbf{0}$.
Output:matrix $B_k \in R^{(k+1)\times k}$, orthonormal bases of $K_k(A^TA, A^Tp_1)$ and $K_{k+1}(AA^T, p_1)$.
for $j = 1, 2, \dots, k$ do

1. $r = A^Tp_j - \beta_j q_{j-1}$;
2. $\alpha_j = \|r\|$; $q_j = r/\alpha_j$;
3. $z = Aq_j - \alpha_j p_j$;
4. $\beta_{j+1} = \|z\|$; $p_{j+1} = z/\beta_{j+1}$;

end for

Let $Q_k = (q_1, q_2, \ldots, q_k)$, $P_{k+1} = (p_1, p_2, \ldots, p_{k+1})$ and

$$B_k = \begin{pmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \beta_3 & \ddots & \\ & & \ddots & \alpha_k \\ & & & \beta_{k+1} \end{pmatrix} \in R^{(k+1)\times k}.$$

It is easy to verify that the following relations hold

$$\begin{aligned} AQ_k &= P_{k+1}B_k, \\ A^T P_{k+1} &= Q_k B_k^T + \alpha_{k+1} q_{k+1} e_{k+1}^T, \end{aligned} \tag{6}$$

where e_{k+1} is the $(k+1)$-th canonical basis vector of R^{k+1}. It follows from (6) that

$$B_k = P_{k+1}^T A Q_k.$$

We assume that the singular values $\theta_i^{(k)} (i = 1, 2, \ldots, k)$ of B_k, called the Ritz values of A with respect to the left and right subspaces $span(P_{k+1}) = K_{k+1}(AA^T, p_1)$ and $span(Q_k) = K_k(A^T A, A^T p_1)$, are all simple and $\alpha_i > 0, \beta_{i+1} > 0$ if the Lanczos bidiagonalization process does not break down until the k-step iteration.

In Algorithm 1 we take $p_1 = Ab/\|Ab\|$, then $\mathcal{P}_k = span(P_k) = K_k(AA^T, Ab)$, $\mathcal{Q}_k = span(Q_k) = K_k(A^T A, A^T Ab)$. Let the initial approximate solution of (1) be $x_0 = 0$. The RRLSQR method determines iterates x_k for $k \geq 1$ such that

$$\|Ax_k - b\| = \min_{x \in span(Q_k)} \|Ax - b\|. \tag{7}$$

Substituting the decomposition (6) and $x = Q_k y$ $(y \in R^k)$ into (7) yields

$$\begin{aligned} \min_{x\in span(Q_k)} \|Ax - b\|^2 &= \min_{y\in R^k} \|AQ_k y - b\|^2 = \min_{y\in R^k} \|P_{k+1}B_k y - b\|^2 \\ &= \min_{y\in R^k} \|B_k y - P_{k+1}^T b\|^2 + b^T (I - P_{k+1}P_{k+1}^T)b. \end{aligned} \tag{8}$$

Since B_k is of full rank, it follows from (8) that the least squares problem $\min_{y\in R^k} \|B_k y - P_{k+1}^T b\|$ has the unique solution $y_k = B_k^\dagger P_{k+1}^T b$. Then the k-th iterate x_k may be expressed as

$$x_k = Q_k y_k = Q_k B_k^\dagger P_{k+1}^T b. \tag{9}$$

With the increase of index k, the residual norm $\|Ax_k - b\|$ decreases and the solution norm $\|x_k\| = \|y_k\|$ increases monotonically, respectively. It follows from (9) that the RRLSQR method solves a sequence of problems

$$\min \|x\| \quad \text{subject to} \quad \|P_{k+1}B_k Q_k^T x - b\| = \min \tag{10}$$

to obtain the regularized solutions x_k of (1) from $k = 1$ upwards. Similar to the TSVD method, which replaces A with the best rank k approximation A_k of A, (10) shows that the RRLSQR iterates can be interpreted as the minimum-norm least squares solutions of the perturbed problems that replace A in (1) by its rank k approximations $P_{k+1}B_k Q_k^T$.

Let

$$\gamma_k = \|A - P_{k+1}B_k Q_k^T\|,$$

which measures the accuracy of the rank k approximation $P_{k+1}B_k Q_k^T$ to A, then we have $\gamma_k \geq \sigma_{k+1}$ since the best rank k approximation A_k satisfies $\|A - A_k\| = \sigma_{k+1}$. If γ_k is closer to σ_{k+1} than σ_k:

$$\sigma_{k+1} \leq \gamma_k < \frac{\sigma_k + \sigma_{k+1}}{2},$$

then $P_{k+1}B_k Q_k^T$ is called a near best rank k approximation to A [33].

The following result can be derived from Property 2.8 in [54].

Lemma 1 *If the RRLSQR method is applied to* (1) *with the starting vector* $p_1 = Ab/\|Ab\|$, *then the* k*-th iterate* x_k *can be expressed as*

$$x_k = \sum_{i=1}^{n} f_i^{(k)} \frac{u_i^T b}{\sigma_i} v_i, \; k = 1, 2, \ldots, n \tag{11}$$

where the filters

$$f_i^{(k)} = 1 - \prod_{j=1}^{k} \frac{(\theta_j^{(k)})^2 - \sigma_i^2}{(\theta_j^{(k)})^2}, \; i = 1, 2, \ldots, n$$

and $\theta_1^{(k)} > \theta_2^{(k)} > \ldots > \theta_k^{(k)}$ *are the singular values of* B_k.

Based on the filtered SVD expansion (11), we obtain the following result concerning the semi-convergence point of the RRLSQR method and the transition point of the TSVD method.

Theorem 1 *Assume that k^* is the semi-convergence point of the RRLSQR method. If the Ritz values $\theta_j^{(k)}$ of B_k converge to the first k^* large singular values σ_j of A in natural order, then $k^* = k_0$; if the Ritz values $\theta_j^{(k)}$ of B_k do not converge to the first k large singular values σ_j of A in natural order for some $k < k^*$, then $k^* < k_0$.*

Proof. The proof is similar to that of Theorem 3.1 in [33]. Hence, it is omitted. □

3. The distances between Krylov subspaces and singular subspaces

In this section, we consider the fundamental problem: how do the underlying k dimensional left and right Krylov subspaces $\mathcal{P}_k = K_k(AA^T, Ab)$ and $\mathcal{Q}_k = K_k(A^TA, A^TAb)$ approximate the k dimensional dominant left and right singular subspaces $\mathcal{U}_k$ and $\mathcal{V}_k$ of A, respectively?

Let

$$D = diag(\sigma_1^2 v_1^T b, \sigma_2^2 v_2^T b, \ldots, \sigma_n^2 v_n^T b) = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}, \tag{12}$$

$$\hat{D} = diag(\sigma_1 v_1^T b, \sigma_2 v_2^T b, \ldots, \sigma_n v_n^T b) = \begin{pmatrix} \hat{D}_1 & 0 \\ 0 & \hat{D}_2 \end{pmatrix}, \tag{13}$$

$$T_k = \begin{pmatrix} 1 & \sigma_1^2 & \ldots & \sigma_1^{2k-2} \\ 1 & \sigma_2^2 & \ldots & \sigma_2^{2k-2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \sigma_n^2 & \ldots & \sigma_n^{2k-2} \end{pmatrix} = \begin{pmatrix} T_{k1} \\ T_{k2} \end{pmatrix}, \tag{14}$$

where $D_1, \hat{D}_1, T_{k1} \in R^{k\times k}$. It follows from (5) and $\sigma_i \neq \sigma_j$ $(i, j = 1, 2, \ldots, k, i \neq j)$ that D_1, $\hat{D}_1$ and T_{k1} are nonsingular. Let

$$\Delta_k = D_2 T_{k2} T_{k1}^{-1} D_1^{-1}, \ \ \hat{\Delta}_k = \hat{D}_2 T_{k2} T_{k1}^{-1} \hat{D}_1^{-1} \in R^{(n-k)\times k}. \tag{15}$$

We present the following $sin\Theta$ theorems which measure the 2-norm distances between $\mathcal{Q}_k$ and $\mathcal{V}_k$, and $\mathcal{P}_k$ and $\mathcal{U}_k$.

Theorem 2 *For $k = 1, 2, \ldots, n-1$, we have*

$$\|sin\Theta(\mathcal{Q}_k, \mathcal{V}_k)\| = \frac{\|\Delta_k\|}{\sqrt{1+\|\Delta_k\|^2}}, \tag{16}$$

$$\|sin\Theta(\mathcal{P}_k, \mathcal{U}_k)\| = \frac{\|\hat{\Delta}_k\|}{\sqrt{1+\|\hat{\Delta}_k\|^2}}, \tag{17}$$

where $\Delta_k, \hat{\Delta}_k \in R^{(n-k)\times k}$ are defined by (15).

Proof. Let $V_n = [v_1, v_2, \ldots, v_n]$, v_i is the i-th right singular vector of A. Using (12) and (14), we obtain $K_k(\Sigma^2, \Sigma^2 V_n^T b) = span(DT_k)$. From $K_k(A^TA, A^TAb) = span(V_n DT_k)$ we have

$$\mathcal{Q}_k = K_k(A^TA, A^TAb) = span\left(V_n \begin{pmatrix} D_1 T_{k1} \\ D_2 T_{k2} \end{pmatrix}\right) = span\left(V_n \begin{pmatrix} I \\ \Delta_k \end{pmatrix}\right).$$

Let $V_n = (V_k, V_k^{\perp})$ with $V_k^{\perp} = [v_{k+1}, v_{k+2}, \ldots, v_n]$, and define

$$Z_k = V_n \begin{pmatrix} I \\ \Delta_k \end{pmatrix} = V_k + V_k^{\perp}\Delta_k, \tag{18}$$

then $Z_k^T Z_k = I + \Delta_k^T \Delta_k$. Let $\hat{Z}_k = Z_k(Z_k^T Z_k)^{-\frac{1}{2}}$, then the columns of $\hat{Z}_k$ form an orthonormal basis of $\mathcal{Q}_k$. From (18) we get

$$\hat{Z}_k = (V_k + V_k^{\perp}\Delta_k)(I + \Delta_k^T \Delta_k)^{-\frac{1}{2}}. \tag{19}$$

It follows from the definition of $\Theta(\mathcal{Q}_k, \mathcal{V}_k)$ and (19) that

$$\|sin\Theta(\mathcal{Q}_k, \mathcal{V}_k)\| = \|(V_k^{\perp})^T \hat{Z}_k\| = \|\Delta_k(I + \Delta_k^T\Delta_k)^{-\frac{1}{2}}\| = \frac{\|\Delta_k\|}{\sqrt{1+\|\Delta_k\|^2}}.$$

Let $U_n = [u_1, u_2, \dots, u_n]$, u_i is the i-th left singular vector of A. Using (13) and (14), we obtain $K_k(\Sigma^2, \Sigma V_n^T b) = span(\hat{D}T_k)$. From $K_k(AA^T, Ab) = span(U_n\hat{D}T_k)$ we have

$$\mathcal{P}_k = K_k(AA^T, Ab) = span\left(U_n\begin{pmatrix}\hat{D}_1T_{k1}\\ \hat{D}_2T_{k2}\end{pmatrix}\right) = span\left(U_n\begin{pmatrix}I\\ \hat{\Delta}_k\end{pmatrix}\right).$$

The remaining proof of (17) is similar to that of (16). Hence, it is omitted. □

Let

$$L_j^{(k)}(\lambda) = \prod_{i=1, i\neq j}^{k} \frac{\lambda - \sigma_i^2}{\sigma_j^2 - \sigma_i^2}, i = 1, 2, \dots, k$$

be the j-th Lagrangian interpolation basis function at nodes $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$, and

$$|L_{k_1}^{(k)}(0)| = \max_{j=1,2,\dots,k} |L_j^{(k)}(0)|, \quad |L_j^{(k)}(0)| = \prod_{i=1, i\neq j}^{k} \frac{\sigma_i^2}{|\sigma_j^2 - \sigma_i^2|}, i = 1, 2, \dots k. \tag{20}$$

$|L_j^{(k)}(0)|(j = 1, 2, \dots, k)$ plays a key role in studying the regularizing effects of the LSQR and RRLSQR methods. Jia [33] provided the accurate estimates of $|L_j^{(k)}(0)|$ for severely, moderately and mildly ill-posed problems.

Lemma 2 *[33] For the severely ill-posed problem with the singular values $\sigma_j = \mathcal{O}(\rho^{-j})$ $(j = 1, 2, \dots n)$ where $\rho > 1$ is an appropriate constant, and $k = 2, 3, \dots, n-1$, we have*

$$|L_k^{(k)}(0)| = 1 + \mathcal{O}(\rho^{-2}),$$

$$|L_j^{(k)}(0)| = \frac{1 + \mathcal{O}(\rho^{-2})}{\prod_{i=j+1}^{k}(\frac{\sigma_j}{\sigma_i})^2} = \frac{1 + \mathcal{O}(\rho^{-2})}{\mathcal{O}(\rho^{(k-j)(k-j+1)})}, \ j = 1, 2, \dots, k-1,$$

$$|L_{k_1}^{(k)}(0)| = \max_{j=1,2,\dots,k} |L_j^{(k)}(0)| = 1 + \mathcal{O}(\rho^{-2}).$$

For a moderately ill-posed problem with singular values $\sigma_j = \zeta j^{-\alpha}$ $(j = 1, 2, \dots, n)$ where $\alpha > 1$ and $\zeta > 0$ are some constants, and $k = 2, 3, \dots, n-1$, we have

$$|L_j^{(k)}(0)| \approx (1 + \frac{j}{2\alpha + 1}) \prod_{i=j+1}^{k} (\frac{j}{i})^{2\alpha}, \ j = 1, 2, \dots, k-1,$$

$$\frac{k}{2\alpha + 1} < |L_{k_1}^{(k)}(0)| \approx 1 + \frac{k}{2\alpha + 1}$$

with the lower bound requiring that k satisfies $\frac{2\alpha+1}{k} \leq 1$. For a mildly ill-posed problem with singular values $\sigma_j = \zeta j^{-\alpha}$ $(j = 1, 2, \dots, n)$ where $\frac{1}{2} < \alpha \leq 1$ and $\zeta > 0$ are some constants, if k satisfies $\frac{2\alpha+1}{k} \leq 1$, we have

$$\frac{k}{2\alpha + 1} < |L_{k_1}^{(k)}(0)|.$$

Now, we establish accurate estimates about $\|\Delta_k\|$ and $\|\hat{\Delta}_k\|$ for severely ill-posed problems.

Theorem 3 *Let the SVD of A be as (3). Assume that (1) is severely ill-posed problem with $\sigma_j = \mathcal{O}(\rho^{-j})$ $(j = 1, 2, \dots, n)$ where $\rho > 1$ is some constant. Then*

$$\|\Delta_1\| \leq \frac{\sigma_2^2}{\sigma_1^2} \frac{max_{2\leq i\leq n}|v_i^T b|}{|v_1^T b|}(1 + \mathcal{O}(\rho^{-4})), \tag{21}$$

$$\|\Delta_k\| \leq \frac{\sigma_{k+1}^2}{\sigma_k^2} \frac{max_{k+1\leq i\leq n}|v_i^T b|}{min_{1\leq i\leq k}|v_i^T b|}(1 + \mathcal{O}(\rho^{-4}))|L_{k_1}^{(k)}(0)|, k = 2, 3, \dots, n-1, \tag{22}$$

$$\|\hat{\Delta}_1\| \leq \frac{\sigma_2}{\sigma_1} \frac{max_{2\leq i\leq n}|v_i^T b|}{|v_1^T b|}(1 + \mathcal{O}(\rho^{-2})), \tag{23}$$

$$\|\hat{\Delta}_k\| \leq \frac{\sigma_{k+1}}{\sigma_k} \frac{max_{k+1\leq i\leq n}|v_i^T b|}{min_{1\leq i\leq k}|v_i^T b|}(1 + \mathcal{O}(\rho^{-2}))|L_{k_1}^{(k)}(0)|, k = 2, 3, \dots, n-1. \tag{24}$$

Proof. By using the structure of the matrix T_k and the property of the Vandermonde matrix T_{k1}, $T_{k2}T_{k1}^{-1}$ may be expressed as [33]

$$T_{k2}T_{k1}^{-1} = \begin{pmatrix} L_1^{(k)}(\sigma_{k+1}^2) & L_2^{(k)}(\sigma_{k+1}^2) & \dots & L_k^{(k)}(\sigma_{k+1}^2) \\ L_1^{(k)}(\sigma_{k+2}^2) & L_2^{(k)}(\sigma_{k+2}^2) & \dots & L_k^{(k)}(\sigma_{k+2}^2) \\ \vdots & \vdots & & \vdots \\ L_1^{(k)}(\sigma_n^2) & L_2^{(k)}(\sigma_n^2) & \dots & L_k^{(k)}(\sigma_n^2) \end{pmatrix} \in R^{(n-k)\times k}. \tag{25}$$

$|L_j^{(k)}(\lambda)|$ is bounded by $|L_j^{(k)}(0)|$ since $|L_j^{(k)}(\lambda)|$ is monotonically decreasing for $0 \le \lambda \le \sigma_k^2$. It follows from this property and the definition (20) of $L_{k_1}^{(k)}(0)$ that $|L_i^{(k)}(\sigma_j^2)| \le |L_{k_1}^{(k)}(0)| (i = 1, 2, \dots, k; j = k+1, k+2, \dots, n)$. From (25) we have

$$\begin{aligned} |\Delta_k| &= |D_2 T_{k2} T_{k1}^{-1} D_1^{-1}| \\ &\le \begin{pmatrix} \frac{\sigma_{k+1}^2}{\sigma_1^2}\left|\frac{v_{k+1}^T b}{v_1^T b}\right| |L_{k_1}^{(k)}(0)| & \frac{\sigma_{k+1}^2}{\sigma_2^2}\left|\frac{v_{k+1}^T b}{v_2^T b}\right| |L_{k_1}^{(k)}(0)| & \dots & \frac{\sigma_{k+1}^2}{\sigma_k^2}\left|\frac{v_{k+1}^T b}{v_k^T b}\right| |L_{k_1}^{(k)}(0)| \\ \frac{\sigma_{k+2}^2}{\sigma_1^2}\left|\frac{v_{k+2}^T b}{v_1^T b}\right| |L_{k_1}^{(k)}(0)| & \frac{\sigma_{k+2}^2}{\sigma_2^2}\left|\frac{v_{k+2}^T b}{v_2^T b}\right| |L_{k_1}^{(k)}(0)| & \dots & \frac{\sigma_{k+2}^2}{\sigma_k^2}\left|\frac{v_{k+2}^T b}{v_k^T b}\right| |L_{k_1}^{(k)}(0)| \\ \vdots & \vdots & & \vdots \\ \frac{\sigma_n^2}{\sigma_1^2}\left|\frac{v_n^T b}{v_1^T b}\right| |L_{k_1}^{(k)}(0)| & \frac{\sigma_n^2}{\sigma_2^2}\left|\frac{v_n^T b}{v_2^T b}\right| |L_{k_1}^{(k)}(0)| & \dots & \frac{\sigma_n^2}{\sigma_k^2}\left|\frac{v_n^T b}{v_k^T b}\right| |L_{k_1}^{(k)}(0)| \end{pmatrix} \\ &= |L_{k_1}^{(k)}(0)| \tilde{\Delta}_k, \end{aligned} \tag{26}$$

where $\tilde{\Delta}_k = (\sigma_{k+1}^2 |v_{k+1}^T b|, \sigma_{k+2}^2 |v_{k+2}^T b|, \dots, \sigma_n^2 |v_n^T b|)^T (\frac{1}{\sigma_1^2 |v_1^T b|}, \frac{1}{\sigma_2^2 |v_2^T b|}, \dots, \frac{1}{\sigma_k^2 |v_k^T b|})$ is a rank one matrix. Using $\|C\| \le \||C|\|$ for any matrix C [51] and (26), we have

$$\|\Delta_k\| \le \||\Delta_k|\| \le |L_{k_1}^{(k)}(0)| \|\tilde{\Delta}_k\| = |L_{k_1}^{(k)}(0)| \left(\sum_{j=k+1}^{n} \sigma_j^4 |v_j^T b|^2\right)^{1/2} \left(\sum_{j=1}^{k} \frac{1}{\sigma_j^4 |v_j^T b|^2}\right)^{1/2}. \tag{27}$$

In the following we estimate the upper bounds of two square root factors in (27) separately.
From $\sigma_j = \mathcal{O}(\rho^{-j})$ $(j = 1, 2, \dots, n)$, and for $k = 1, 2, \dots, n-1$, we get

$$\begin{aligned} (\sum_{j=k+1}^{n} \sigma_j^4 |v_j^T b|^2)^{1/2} &= \sigma_{k+1}^2 \max_{k+1 \le i \le n} |v_i^T b| \left(\sum_{j=k+1}^{n} \frac{\sigma_j^4 |v_j^T b|^2}{\sigma_{k+1}^4 \max_{k+1\le i\le n} |v_i^T b|^2}\right)^{1/2} \\ &\le \sigma_{k+1}^2 \max_{k+1 \le i \le n} |v_i^T b| \left(\sum_{j=k+1}^{n} \frac{\sigma_j^4}{\sigma_{k+1}^4}\right)^{1/2} \\ &= \sigma_{k+1}^2 \max_{k+1 \le i \le n} |v_i^T b| \left(1 + \sum_{j=k+2}^{n} \mathcal{O}(\rho^{4(k-j)+4})\right)^{1/2} \\ &= \sigma_{k+1}^2 \max_{k+1 \le i \le n} |v_i^T b| \left(1 + \mathcal{O}(\sum_{j=k+2}^{n} \rho^{4(k-j)+4})\right)^{1/2} \\ &= \sigma_{k+1}^2 \max_{k+1 \le i \le n} |v_i^T b| \left(1 + \mathcal{O}[\frac{\rho^{-4}}{1-\rho^{-4}}(1-\rho^{-4(n-k-1)})]\right)^{1/2} \\ &= \sigma_{k+1}^2 \max_{k+1 \le i \le n} |v_i^T b| (1 + \mathcal{O}(\rho^{-4}))^{1/2} \\ &= \sigma_{k+1}^2 \max_{k+1 \le i \le n} |v_i^T b| (1 + \mathcal{O}(\rho^{-4})) \end{aligned} \tag{28}$$

with $1+\mathcal{O}(\rho^{-4})=1$ for $k=n-1$. For $k=2,3,\ldots,n-1$, we obtain

$$
\begin{aligned}
(\sum_{j=1}^{k}\frac{1}{\sigma_j^4|v_j^Tb|^2})^{1/2} &= \frac{1}{\sigma_k^2\min_{1\le i\le k}|v_i^Tb|}\left(\sum_{j=1}^{k}\frac{\sigma_k^4\min_{1\le i\le k}|v_i^Tb|^2}{\sigma_j^4|v_j^Tb|^2}\right)^{1/2}\\
&\le \frac{1}{\sigma_k^2\min_{1\le i\le k}|v_i^Tb|}\left(\sum_{j=1}^{k}\frac{\sigma_k^4}{\sigma_j^4}\right)^{1/2}\\
&= \frac{1}{\sigma_k^2\min_{1\le i\le k}|v_i^Tb|}\left(1+\mathcal{O}(\sum_{j=1}^{k-1}\rho^{4(j-k)})\right)^{1/2}\\
&= \frac{1}{\sigma_k^2\min_{1\le i\le k}|v_i^Tb|}(1+\mathcal{O}(\rho^{-4}))^{1/2}\\
&= \frac{1}{\sigma_k^2\min_{1\le i\le k}|v_i^Tb|}(1+\mathcal{O}(\rho^{-4})).
\end{aligned}
\tag{29}
$$

From (27)-(29), we obtain (22).

For the special case of $k=1$, from (12) and (14) we have

$$
D_1=\sigma_1^2v_1^Tb,\ \ D_2=diag(\sigma_2^2v_2^Tb,\sigma_3^2v_3^Tb,\ldots,\sigma_n^2v_n^Tb),
$$

$$
T_{k1}=1,\ \ T_{k2}=(1,1,\ldots,1)^T\in R^{n-1}.
$$

Using (15), we obtain

$$
\Delta_1=\frac{1}{\sigma_1^2v_1^Tb}(\sigma_2^2v_2^Tb,\sigma_3^2v_3^Tb,\ldots,\sigma_n^2v_n^Tb)^T. \tag{30}
$$

From (28) and (30) we get (21).

Using (13), (15) and (25), similar to (27), we can obtain

$$
\|\hat{\Delta}_k\|\le|L_{k_1}^{(k)}(0)|\left(\sum_{j=k+1}^{n}\sigma_j^2|v_j^Tb|^2\right)^{1/2}\left(\sum_{j=1}^{k}\frac{1}{\sigma_j^2|v_j^Tb|^2}\right)^{1/2}. \tag{31}
$$

Similar to the derivation of the inequalities (28) - (30), we can prove both (23) and (24) by using (31). □

It follows from Lemma 2 and Theorem 3 that the bounds (21) - (24) can be unified as

$$
\|\Delta_k\|\le\frac{\sigma_{k+1}^2}{\sigma_k^2}\frac{max_{k+1\le i\le n}|v_i^Tb|}{min_{1\le i\le k}|v_i^Tb|}(1+\mathcal{O}(\rho^{-2})),\ \ k=1,2,\ldots,n-1, \tag{32}
$$

$$
\|\hat{\Delta}_k\|\le\frac{\sigma_{k+1}}{\sigma_k}\frac{max_{k+1\le i\le n}|v_i^Tb|}{min_{1\le i\le k}|v_i^Tb|}(1+\mathcal{O}(\rho^{-2})),\ \ k=1,2,\ldots,n-1. \tag{33}
$$

Remark 1 *Based on (4),(5) and $\mathcal{E}(|v_i^Te|)=\eta$ $(i=1,2,\ldots,n)$, we have*

$$
\frac{max_{k+1\le i\le n}|v_i^Tb|}{min_{1\le i\le k}|v_i^Tb|}\approx\frac{max_{k+1\le i\le n}|v_i^Tb_{true}|}{min_{1\le i\le k}|v_i^Tb_{true}|}\approx\frac{|v_{k+1}^Tb_{true}|}{|v_k^Tb_{true}|}\approx\frac{\sigma_{k+1}^{1+\beta}}{\sigma_k^{1+\beta}}<1,\ \ k=1,2,\ldots,k_0-1, \tag{34}
$$

$$
\frac{max_{k_0+1\le i\le n}|v_i^Tb|}{min_{1\le i\le k_0}|v_i^Tb|}\approx\frac{max_{k_0+1\le i\le n}|v_i^Te|}{min_{1\le i\le k_0}|v_i^Tb_{true}|}\approx\frac{\eta}{\sigma_{k_0}^{1+\beta}}<1, \tag{35}
$$

$$
\frac{max_{k+1\le i\le n}|v_i^Tb|}{min_{1\le i\le k}|v_i^Tb|}\approx\frac{max_{k_0+1\le i\le n}|v_i^Te|}{min_{1\le i\le k_0}|v_i^Te|}\approx\frac{\eta}{\eta}\approx1,\ \ k=k_0+1,\ldots,n-1. \tag{36}
$$

Remark 2 *For severely ill-posed problems, from (32) - (36), we have*

$$
\|\Delta_k\|\le\frac{\sigma_{k+1}^2}{\sigma_k^2}\cdot\frac{\sigma_{k+1}^{1+\beta}}{\sigma_k^{1+\beta}}(1+\mathcal{O}(\rho^{-2}))\sim\rho^{-(3+\beta)},\ \ k=1,2,\ldots,k_0, \tag{37}
$$

$$
\|\Delta_k\|\le\frac{\sigma_{k+1}^2}{\sigma_k^2}(1+\mathcal{O}(\rho^{-2}))\sim\rho^{-2},\ \ k=k_0+1,\ldots,n-1, \tag{38}
$$

$$
\|\hat{\Delta}_k\|\le\frac{\sigma_{k+1}}{\sigma_k}\cdot\frac{\sigma_{k+1}^{1+\beta}}{\sigma_k^{1+\beta}}(1+\mathcal{O}(\rho^{-2}))\sim\rho^{-(2+\beta)},\ \ k=1,2,\ldots,k_0, \tag{39}
$$

$$\|\hat{\Delta}_k\| \leq \frac{\sigma_{k+1}}{\sigma_k}(1+\mathcal{O}(\rho^{-2})) \sim \rho^{-1},\ \ k = k_0+1,\ldots,n-1. \tag{40}$$

From Theorem 2 and (37) - (40) we observe that both $sin\Theta(\mathcal{Q}_k, \mathcal{V}_k)$ and $sin\Theta(\mathcal{P}_k, \mathcal{U}_k)$ in the RRLSQR method exhibit neither increasing nor decreasing tendency for $k = 1,2,\ldots,k_0$ and $k = k_0+1,\ldots,n-1$, respectively. For the LSQR method, Jia [33] showed that

$$\|\Delta_k\| \leq \frac{\sigma_{k+1}}{\sigma_k}\cdot\frac{\sigma_{k+1}^{1+\beta}}{\sigma_k^{1+\beta}}(1+\mathcal{O}(\rho^{-2})) \sim \rho^{-(2+\beta)},\ \ k = 1,2,\ldots,k_0, \tag{41}$$

$$\|\Delta_k\| \leq \frac{\sigma_{k+1}}{\sigma_k}(1+\mathcal{O}(\rho^{-2})) \sim \rho^{-1},\ \ k = k_0+1,\ldots,n-1. \tag{42}$$

(37), (38) and (41), (42) show that the upper bound of $\|\Delta_k\|$ in the RRLSQR method is smaller than that of $\|\Delta_k\|$ in the LSQR method. It follows from Theorem 2, (37), (38), (41) and (42) that the upper bound of $sin\Theta(\mathcal{Q}_k, \mathcal{V}_k)$ in the RRLSQR method is smaller than that of $sin\Theta(\mathcal{V}_k^R, \mathcal{V}_k)$ in the LSQR method, where $\mathcal{V}_k^R = K_k(A^TA, A^Tb)$, which means that the Krylov subspace $\mathcal{Q}_k$ captures more information about $\mathcal{V}_k$ than $\mathcal{V}_k^R$.

Next we estimate $\|\Delta_k\|$ and $\|\hat{\Delta}_k\|$ for moderately and mildly ill-posed problems.

Theorem 4 *For a moderately or mildly ill-posed problem with singular values $\sigma_j = \zeta j^{-\alpha}$ $(j = 1,2,\ldots,n)$ where $\zeta > 0$ and $\alpha > 1$ or $\frac{1}{2} < \alpha \leq 1$, we obtain*

$$\|\Delta_1\| \leq \frac{max_{2\leq i\leq n}|v_i^Tb|}{|v_1^Tb|}\frac{1}{\sqrt{4\alpha-1}}, \tag{43}$$

$$\|\Delta_k\| \leq \frac{max_{k+1\leq i\leq n}|v_i^Tb|}{min_{1\leq i\leq k}|v_i^Tb|}\sqrt{\frac{k^2}{16\alpha^2-1}+\frac{k}{4\alpha-1}}|L_{k_1}^{(k)}(0)|,\ \ k = 2,3,\ldots,n-1, \tag{44}$$

$$\|\hat{\Delta}_1\| \leq \frac{max_{2\leq i\leq n}|v_i^Tb|}{|v_1^Tb|}\frac{1}{\sqrt{2\alpha-1}}, \tag{45}$$

$$\|\hat{\Delta}_k\| \leq \frac{max_{k+1\leq i\leq n}|v_i^Tb|}{min_{1\leq i\leq k}|v_i^Tb|}\sqrt{\frac{k^2}{4\alpha^2-1}+\frac{k}{2\alpha-1}}|L_{k_1}^{(k)}(0)|,\ \ k = 2,3,\ldots,n-1. \tag{46}$$

Proof. We only need to estimate the right side of (27). For $k = 1,2,\ldots,n-1$, we have

$$\begin{aligned}
(\sum_{j=k+1}^{n}\sigma_j^4|v_j^Tb|^2)^{1/2} &= \sigma_{k+1}^2\max_{k+1\leq i\leq n}|v_i^Tb|\left(\sum_{j=k+1}^{n}\frac{\sigma_j^4|v_j^Tb|^2}{\sigma_{k+1}^4\max_{k+1\leq i\leq n}|v_i^Tb|^2}\right)^{1/2}\\
&\leq \sigma_{k+1}^2\max_{k+1\leq i\leq n}|v_i^Tb|\left(\sum_{j=k+1}^{n}(\frac{j}{k+1})^{(-4\alpha)}\right)^{1/2}\\
&= \sigma_{k+1}^2\max_{k+1\leq i\leq n}|v_i^Tb|\left((k+1)^{4\alpha}\sum_{j=k+1}^{n}\frac{1}{j^{4\alpha}}\right)^{1/2}\\
&\leq \sigma_{k+1}^2\max_{k+1\leq i\leq n}|v_i^Tb|(k+1)^{2\alpha}\left(\int_k^\infty\frac{1}{x^{4\alpha}}dx\right)^{1/2}\\
&= \sigma_{k+1}^2\max_{k+1\leq i\leq n}|v_i^Tb|(\frac{\sigma_k}{\sigma_{k+1}})^2\sqrt{\frac{k}{4\alpha-1}}\\
&= \sigma_k^2\max_{k+1\leq i\leq n}|v_i^Tb|\sqrt{\frac{k}{4\alpha-1}}.
\end{aligned} \tag{47}$$

For $\alpha > 1$ or $\frac{1}{2} < \alpha \le 1$ the function $x^{4\alpha}$ is convex over the interval $[0, 1]$, we get

$$\begin{aligned}(\sum_{j=1}^{k} \frac{1}{\sigma_j^4 |v_j^T b|^2})^{1/2} &= \frac{1}{\sigma_k^2 \min_{1\le i\le k} |v_i^T b|} \left(\sum_{j=1}^{k} \frac{\sigma_k^4 \min_{1\le i\le k} |v_i^T b|^2}{\sigma_j^4 |v_j^T b|^2} \right)^{1/2} \\ &\le \frac{1}{\sigma_k^2 \min_{1\le i\le k} |v_i^T b|} \left(\sum_{j=1}^{k} (\frac{j}{k})^{4\alpha} \right)^{1/2} \\ &= \frac{1}{\sigma_k^2 \min_{1\le i\le k} |v_i^T b|} \left(k \sum_{j=1}^{k} \frac{1}{k} (\frac{j-1}{k})^{4\alpha} + 1 \right)^{1/2} \\ &\le \frac{1}{\sigma_k^2 \min_{1\le i\le k} |v_i^T b|} \left(k \int_0^1 x^{4\alpha} dx + 1 \right)^{1/2} \\ &\le \frac{1}{\sigma_k^2 \min_{1\le i\le k} |v_i^T b|} \sqrt{\frac{k}{1+4\alpha} + 1}, \ \ k = 2, 3, \ldots, n-1.\end{aligned} \tag{48}$$

Substituting (47) and (48) into (27) yields (44). For $k = 1$, (43) follows from (30) and (47).

Using (31), we can prove both (45) and (46), similar to Theorem 4.4 in [33]. Hence, it is omitted. □

Remark 3 *For moderately and mildly ill-posed problems and the RRLSQR method, based on (34) - (36) and (43)- (46), we have*

$$\|\Delta_1\| \le \frac{\sigma_2^{1+\beta}}{\sigma_1^{1+\beta}} \frac{1}{\sqrt{4\alpha - 1}} \sim 2^{-\alpha(1+\beta)} \frac{1}{\sqrt{4\alpha - 1}}, \tag{49}$$

$$\begin{aligned}\|\Delta_k\| &\le \frac{\sigma_{k+1}^{1+\beta}}{\sigma_k^{1+\beta}} \sqrt{\frac{k^2}{16\alpha^2 - 1} + \frac{k}{4\alpha - 1}} |L_{k_1}^{(k)}(0)| \\ &\sim (\frac{k}{k+1})^{\alpha(1+\beta)} \sqrt{\frac{k^2}{16\alpha^2 - 1} + \frac{k}{4\alpha - 1}} |L_{k_1}^{(k)}(0)|, \ \ k = 2, 3, \ldots, n-1\end{aligned} \tag{50}$$

and

$$\begin{aligned}\|\hat{\Delta}_1\| &\le \frac{\sigma_2^{1+\beta}}{\sigma_1^{1+\beta}} \frac{1}{\sqrt{2\alpha - 1}} \sim 2^{-\alpha(1+\beta)} \frac{1}{\sqrt{2\alpha - 1}}, \\ \|\hat{\Delta}_k\| &\le \frac{\sigma_{k+1}^{1+\beta}}{\sigma_k^{1+\beta}} \sqrt{\frac{k^2}{4\alpha^2 - 1} + \frac{k}{2\alpha - 1}} |L_{k_1}^{(k)}(0)| \\ &\sim (\frac{k}{k+1})^{\alpha(1+\beta)} \sqrt{\frac{k^2}{4\alpha^2 - 1} + \frac{k}{2\alpha - 1}} |L_{k_1}^{(k)}(0)|, \ \ k = 2, 3, \ldots, n-1.\end{aligned} \tag{51}$$

For the LSQR method, Jia [33] presented

$$\|\Delta_1\| \le \frac{\sigma_2^{1+\beta}}{\sigma_1^{1+\beta}} \frac{1}{\sqrt{2\alpha - 1}} \sim 2^{-\alpha(1+\beta)} \frac{1}{\sqrt{2\alpha - 1}}, \tag{52}$$

and

$$\begin{aligned}\|\Delta_k\| &\le \frac{\sigma_{k+1}^{1+\beta}}{\sigma_k^{1+\beta}} \sqrt{\frac{k^2}{4\alpha^2 - 1} + \frac{k}{2\alpha - 1}} |L_{k_1}^{(k)}(0)| \\ &\sim (\frac{k}{k+1})^{\alpha(1+\beta)} \sqrt{\frac{k^2}{4\alpha^2 - 1} + \frac{k}{2\alpha - 1}} |L_{k_1}^{(k)}(0)|, \ \ k = 2, 3, \ldots, n-1.\end{aligned} \tag{53}$$

Comparing (49) and (50) with (52) and (53), we know that the upper bound of $\|\Delta_k\|$ in RRLSQR method is smaller than that of $\|\Delta_k\|$ in the LSQR method for moderately and mildly ill-posed problems, then the upper bound of $sin\Theta(\mathcal{Q}_k, \mathcal{V}_k)$ is smaller than that of $sin\Theta(\mathcal{V}_k^R, \mathcal{V}_k)$.

Lemma 2 shows that both $\sqrt{\frac{k^2}{16\alpha^2-1} + \frac{k}{4\alpha-1}} |L_{k_1}^{(k)}(0)|$ and $\sqrt{\frac{k^2}{4\alpha^2-1} + \frac{k}{2\alpha-1}} |L_{k_1}^{(k)}(0)|$ increase as k grows. For moderately ill-posed problems, from Theorem 2, (50) and (51) we observe that both $sin\Theta(\mathcal{Q}_k, \mathcal{V}_k)$ and $sin\Theta(\mathcal{P}_k, \mathcal{U}_k)$ in the RRLSQR method exhibit increasing tendency with k, which means that $\mathcal{Q}_k$ and $\mathcal{P}_k$ can not capture $\mathcal{V}_k$ and $\mathcal{U}_k$, respectively. In fact, both $sin\Theta(\mathcal{Q}_k, \mathcal{V}_k)$ and $sin\Theta(\mathcal{P}_k, \mathcal{U}_k)$ approach one as k increases, which means that $\mathcal{Q}_k$ and $\mathcal{P}_k$ will contain substantial information on the right and left singular vectors corresponding to the $n - k$ small singular values of A, respectively. For mildly ill-posed problems, Lemma 2 indicates that $|L_{k_1}^{(k)}(0)|$ is substantially greater than one for $\frac{1}{2} < \alpha \le 1$. Consequently, the upper bounds (50) and (51) become increasingly large as k increases, causing that both $\|\Delta_k\|$ and $\|\hat{\Delta}_k\|$ are large and both $sin\Theta(\mathcal{Q}_k, \mathcal{V}_k)$ and $sin\Theta(\mathcal{P}_k, \mathcal{U}_k)$ approach one.

For the LSQR method, Jia [33] investigated how $\|sin\Theta(\mathcal{V}_k^R, \mathcal{V}_k)\|$ affects the smallest Ritz value $\theta_k^{(k)}$, and showed that $\theta_k^{(k)} > \sigma_{k+1}$ holds for suitable $\|sin\Theta(\mathcal{V}_k^R, \mathcal{V}_k)\| < 1$ and the k Ritz values $\theta_i^{(k)}$ do not approximate the large singular values σ_i of A in natural order when $\|sin\Theta(\mathcal{V}_k^R, \mathcal{V}_k)\|$ is sufficiently close to one. The same results can be obtained for the RRLSQR method, and will be confirmed numerically later.

4. A regularized RRLSQR method

In this section, some computational issues related to the application of the RRLSQR method to the solution of large-scale discrete ill-posed problems are considered, and a regularized RRLSQR method is developed for solving the ill-posed problem (1).

To solve the large-scale linear discrete ill-posed problem (1) by the RRLSQR method, we may compute the solution y_k of a sequence of the least squares problems (8) for increasing iteration number k, and obtain an approximate solution x_k by (9). This is a regularization method since the ill-conditioned problem (1) is replaced by the less ill-conditioned least squares problem (8). The iteration number k plays the role of the regularization parameter determined by the following Stopping Rule based on the discrepancy principle [25].

Stopping Rule 1. *Let $\alpha > 0$ be fixed and let b is the right hand of* (1) *contaminated by errors. Denote the solutions computed by the RRLSQR method applied to the solution of the linear system* (1) *by $x_k(k = 1, 2, \ldots)$. Terminate the RRLSQR iterations as soon as a determined iterate x_k satisfies*

$$\|b - Ax_k\| \leq \alpha\delta.$$

The termination index is denoted by k^*.

However, the condition numbers of the least squares problems (10) are increasing with the increasing iteration number k. Therefore, it is necessary to apply regularization method to solve the least squares problem (10). The TSVD method [25, 22, 24] is a good choice to compute the regularized solution to the least squares problem (10) since the size of the projected problem (10) is not large.

Let

$$B_k = U^{(k)}\Sigma^{(k)}(V^{(k)})^T,$$

be a SVD of the matrix B_k, where $U^{(k)} = [u_1^{(k)}, u_2^{(k)}, \ldots, u_k^{(k)}] \in R^{(k+1)\times k}$ and $V^{(k)} = [v_1^{(k)}, v_2^{(k)}, \ldots, v_k^{(k)}] \in R^{k\times k}$ are orthonormal, $\Sigma_k = diag(\sigma_1^{(k)}, \sigma_2^{(k)}, \ldots, \sigma_k^{(k)})$ with the singular values $\sigma_1^{(k)} \geq \sigma_2^{(k)} \geq \ldots \geq \sigma_k^{(k)} \geq 0$. Then the TSVD regularization solution of the least squares problem (10) computed by the TSVD method can be described as

$$y_k = \sum_{i=1}^{\tilde{k}} \frac{(u_i^{(k)})^T P_{k+1}^T b}{\sigma_i^{(k)}} v_i^{(k)}, \tag{54}$$

where the index $\tilde{k}$ is the regularization parameter, which determined by the generalized cross-validation (GCV) [25, 19]. It determines how many large SVD components of B_k are used to compute a regularization solution y_k to the least squares problem (10). Based on them, an approximate solution to the large-scale linear discrete ill-posed problem (1) can be expressed as

$$x_k = Q_k y_k.$$

If x_k satisfies the Stopping Rule 1, then x_k can be considered as the regularized solution to the ill-posed problem (1). The regularized RRLSQR method for solving the large-scale discrete ill-posed problem (1) can be described as follows.

Algorithm 2. Regularized RRLSQR algorithm.

Input: $A \in R^{n\times n}$, $b \in R^n$.
Output: Regularized solution x_{k^*} of the linear discrete ill-posed problem (1).

1. Compute $p_1 = Ab/\|Ab\|$, and set $k = 1$.
2. Construct the orthonormal basis $\{p_i\}_{i=1}^{k+1}$ of $K_k(AA^T, Ab)$ and $\{q_i\}_{i=1}^{k}$ of $K_k(A^TA, A^TAb)$ by using Algorithm 1.
3. Use the TSVD method to compute the regularized solution (54) of the projected problem (10), and compute the approximate solution $x_k = Q_k x_k$ of (1).
4. If x_k satisfies the Stopping Rule 1, then x_k is the regularized solution of (1), and exit; else set $k := k + 1$ and go to 2.

5. Experimental results

In this section, we will conduct some experiments to confirm our theoretical results and demonstrate the effectiveness of the regularized RRLSQR method for solving large-scale linear discrete ill-posed problems. All the computations were carried out in

MATLAB R2016b on personal computer (3.4 GHz Intel Core i5, 8 GB 1600 MHZ DDR3) with double precision. We choose three ill-posed problems: the severely ill-posed problem $shaw$[24], the moderately ill-posed problem $deriv2$[24] and the mildly ill-posed problem $regutm$[24]. In all test problems, we have not used any preconditioning since preconditioning would change the condition numbers of these problems.

For all the experiments, the exact solution $x_{true} = rand(n,1)$, where function rand creates an $n \times 1$ random matrix with entries uniformly distributed in $[0,1]$, and the error-free right-hand side is $b_{true} = Ax_{true}$. The error is $e = \delta \times \hat{e}/\|\hat{e}\|$, $\hat{e} = rand(n,1)$ but different from x_{true}, $\delta = 10^{-2}, 10^{-3}$. Then the error-contamination right-hand side $b = b_{true} + e$.

For each test problem with $\delta = 10^{-2}$ and $\delta = 10^{-3}$ we give a comparison between $sin\Theta_1 = \|sin\Theta(\mathcal{V}_k^R, \mathcal{V}_k)\|$ and $sin\Theta_2 = \|sin\Theta(\mathcal{Q}_k, \mathcal{V}_k)\|$, where $\Theta(\mathcal{V}_k^R, \mathcal{V}_k)$ and $\Theta(\mathcal{Q}_k, \mathcal{V}_k)$ are the actually canonical angles between subspaces $\mathcal{V}_k^R = K_k(A^TA, A^Tb)$ and $\mathcal{V}_k$ spanned by the k dominant right singular vectors of A, and between subspaces $\mathcal{Q}_k = K_k(A^TA, A^TAb)$ and $\mathcal{V}_k$, respectively, and a comparison between the k Ritz values $\theta_i^{(k)}$ generated by the RRLSQR method and the first $k+1$ large singular values of A. We depict $sin\Theta_1$ versus $sin\Theta_2$ in Figs. 1-6(a) and draw the comparison diagrams of k Ritz values $\theta_i^{(k)}$ and first $k+1$ large singular values σ_i of A which are converted into logarithmic values in Figs. 1-6(b). We also compare the accuracy of the solution x_k^{rrlsqr} obtained by the regularized RRLSQR method with that of the solution x_k^{lsqr} computed by the regularized LSQR method by the relative errors defined as $\frac{\|x_k^{rrlsqr} - x_{true}\|}{\|x_{true}\|}$ and $\frac{\|x_k^{lsqr} - x_{true}\|}{\|x_{true}\|}$. The logarithmic values of the relative errors are shown in Figs. 1-6(c).

Example 1 *Consider the severely ill-posed problem* $A = shaw(1000)$*[24] with* $\delta = 10^{-2}$ *and* $\delta = 10^{-3}$*. Figs. 1-2(a) show that* $sin\Theta_1$ *is greater than or equal to* $sin\Theta_2$ *at most points for* $k = 1, 2, \ldots, 20$*, and indicate that the Krylov subspace* $K_k(A^TA, A^TAb)$ *approximates the k dimensional dominant right singular subspace* V_k *of A more accurately than the Krylov subspace* $K_k(A^TA, A^Tb)$ *in most cases, which justifies the Remark 2. Figs. 1-2(b) show that the smallest Ritz values* $\theta_k^{(k)}$ *of* B_k *are above* σ_{k+1} *for* $k = 1, 2, \ldots, 20$*. Figs. 1-2(c) show that the solution* x_k^{rrlaqr} *obtained by the regularized RRLSQR method has nearly the same precision as the solution* x_k^{lsqr} *found by the LSQR method. In addition, in order to verify the rationality of (5), we compare the* $log|v_k^Tb|, log|v_k^Tb_{true}|$ *with* $log|v_k^Te|$ *for* $A = shaw(1000)$ *with* $\delta = 10^{-2}$*. The transition point* $k_0 = 11$ *of the regularization parameter of the TSVD method is determined by GCV, which satisfies* $\|x_{k_0}^{tsvd} - x_{true}\| = \min_{1 \le k \le n} \|x_k^{tsvd} - x_{true}\|$*. Fig. 1(d) shows that the inequality (5) is reasonable, and the solution* x_k^{rrlaqr} *obtained by the regularized RRLSQR method has nearly the same precision as* $x_{k_0}^{tsvd}$ *determined by the TSVD method.*

For the large-scale ill-posed problems, the TSVD method is time-consuming, and won't even work due to storage requirements. For example, for the ill-posed problem $A = shaw(10000)$ with $\delta = 10^{-3}$, the regularized RRLSQR method only needs 87.92s to get the best regularized solution, while the TSVD method requires 276.45s.

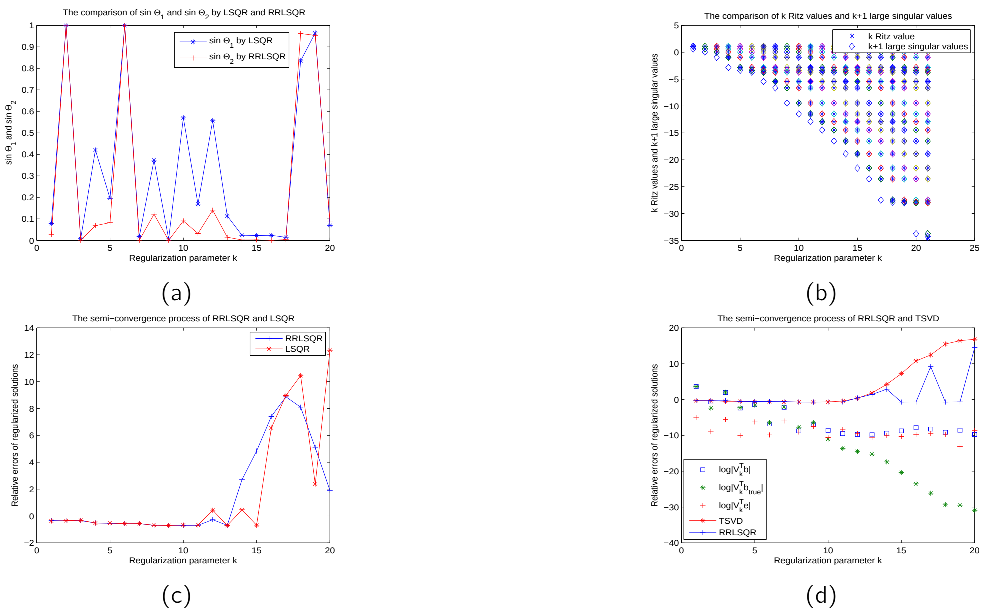
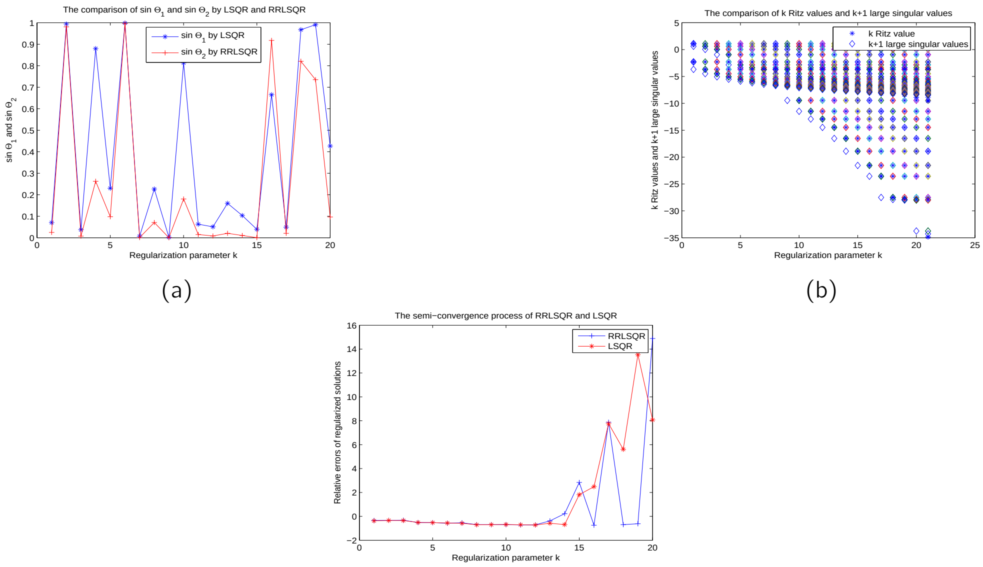


Figure 1. (a) The comparison between $sin\Theta_1$ and $sin\Theta_2$;(b) k Ritz values and the first $k+1$ large singular values of shaw;(c) The relative errors of RRLSQR and LSQR; (d) The relative errors of RRLSQR and TSVD

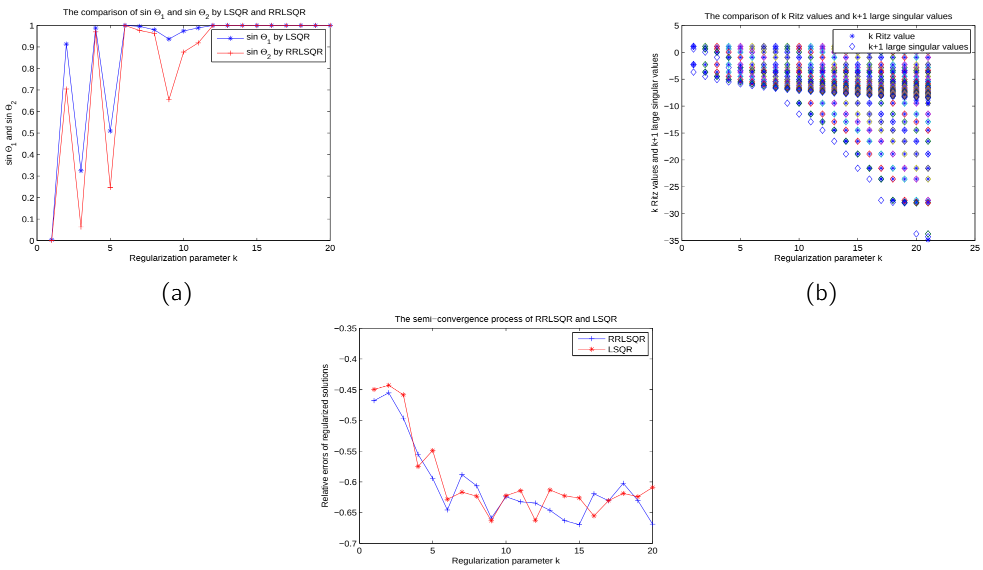
Example 2 *Consider the moderately ill-posed problem* $A = deriv2(1000)$*[24] with* $\delta = 10^{-2}$ *and* $\delta = 10^{-3}$*. Figs. 3-4(a) show that* $sin\Theta_1$ *are also greater than or equal to* $sin\Theta_2$ *for* $k = 1, 2, \ldots, 20$*, which means that the Krylov subspace* $\mathcal{Q}_k$ *captures more information about* $\mathcal{V}_k$ *than* $\mathcal{V}_k^R$*. From Figs. 3-4(a) we observe that from some k upwards, both* $sin\Theta_1$ *and* $sin\Theta_2$ *almost*



(c)

Figure 2. (a) The comparison between $sin\Theta_1$ and $sin\Theta_2$;(b) k Ritz values and the first $k+1$ large singular values of shaw;(c) The relative errors of RRLSQR and LSQR

equal to 1, which justifies the Remark 3. Figs. 3-4(b) show that the k Ritz values of B_k fail to approximate the large singular values of A in natural order. The solution x_k^{rrlaqr} obtained by the regularized RRLSQR method has nearly the same precision as x_k^{lsqr} determined by the regularized LSQR method as shown by Figs. 3-4(c).



(c)

Figure 3. (a) The comparison between $sin\Theta_1$ and $sin\Theta_2$;(b) k Ritz values and the first $k+1$ large singular values of deriv2;(c) The relative errors of RRLSQR and LSQR

Example 3 *Consider the mildly ill-posed problem $A = regutm(1000)$[24] with $\delta = 10^{-2}$ and $\delta = 10^{-3}$. Figs. 5-6(a) indicate that both $sin\Theta_1$ and $sin\Theta_2$ exhibit monotonically increasing tendency, and approach one as k increases faster than both $sin\Theta_1$ and $sin\Theta_2$ in Figs. 3-4(a) because of $\frac{1}{2} < \alpha \leq 1$, which confirms our results. Figs. 5-6(b) show that the smallest Ritz value $\theta_k^{(k)} < \sigma_{k+1}$ from $k = 11$ onwards. Figs. 5-6(c) indicate that the solution x_k^{rrlaqr} obtained by the regularized RRLSQR method has nearly the same precision as x_k^{lsqr} computed by the regularized LSQR method.*

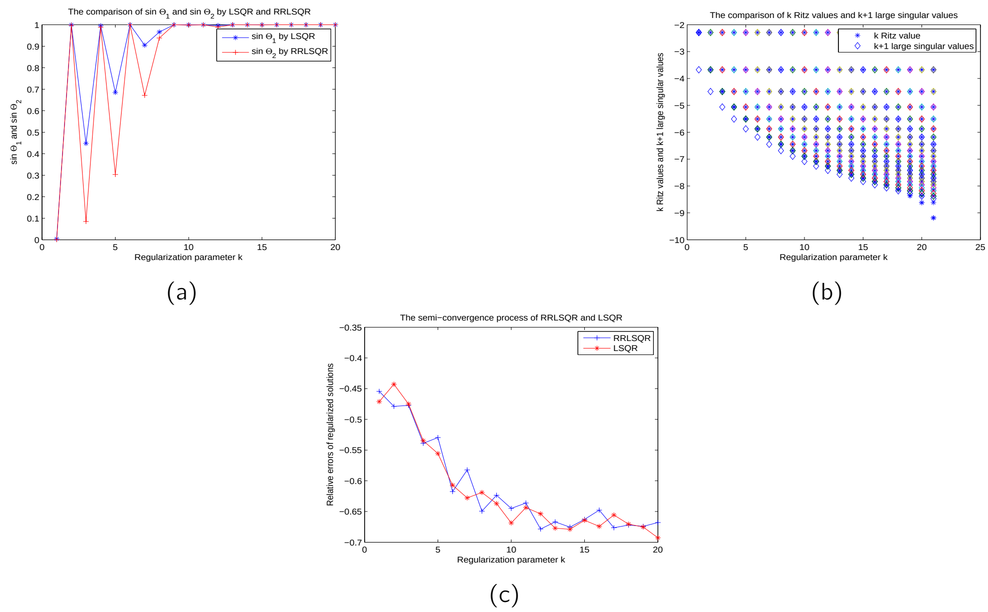


Figure 4. (a) The comparison between $sin\Theta_1$ and $sin\Theta_2$;(b) k Ritz values and the first $k+1$ large singular values of deriv2;(c) The relative errors of RRLSQR and LSQR

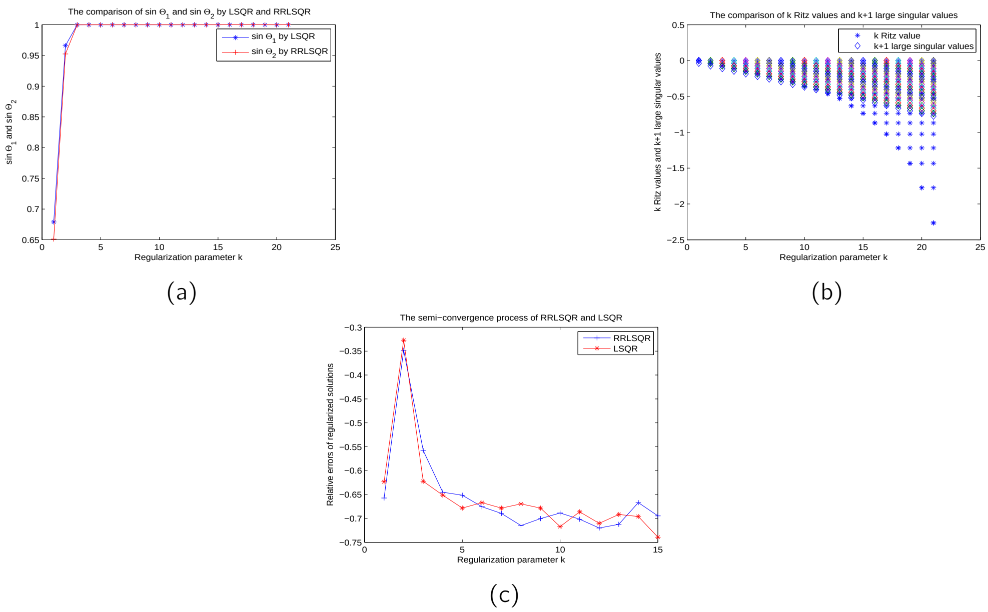


Figure 5. The comparison between $sin\Theta_1$ and $sin\Theta_2$;(b) k Ritz values and the first $k+1$ large singular values of regutm;(c) The relative errors of RRLSQR and LSQR

Example 4 *Consider the ill-posed problem $Ax = b$, $A = regutm(n)$[24] with $\delta = 10^{-2}$, $n = 200, 400, 800, 1000$. The coefficient matrix A is replaced by $A - \sigma_{n/10}I$. The resulting problem $(A - \sigma_{n/10}I)x = b$ is inconsistent. We apply the regularized LSQR method and the regularized RRLSQR method to solve the inconsistent linear system. The numerical results are in Table 1.*

From Table 1 we observe that the regularized RRLSQR method gives more accurate computed solutions than the regularized LSQR method.

6. Conclusions

For large-scale linear discrete ill-posed problems, iterative solvers have received considerable attention over the decades. Of them, the Krylov iterative solver LSQR is most widely used. In this paper, we develop a variant of the LSQR method, the

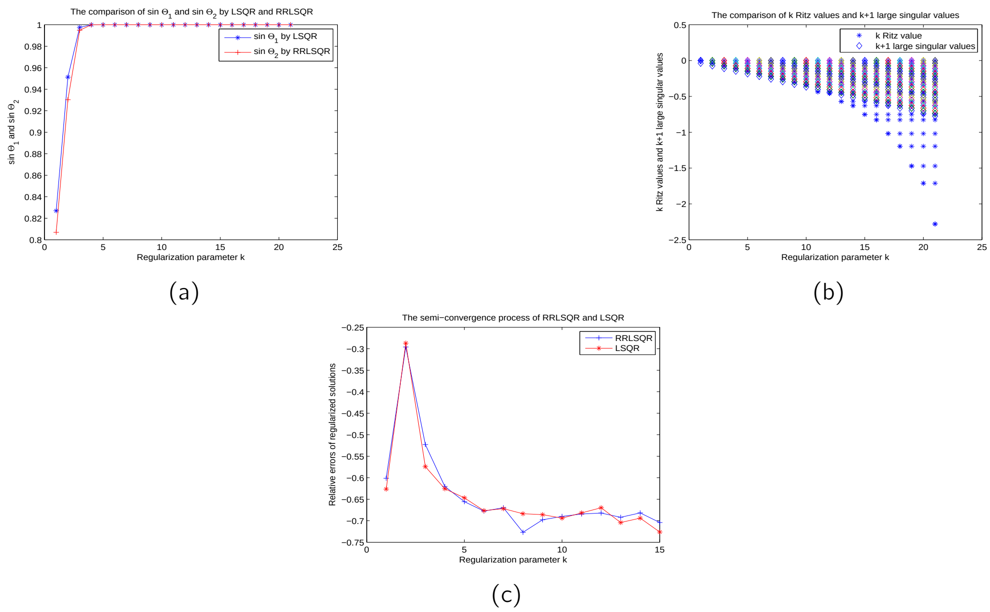


Figure 6. The comparison between $sin\Theta_1$ and $sin\Theta_2$;(b) k Ritz values and the first $k+1$ large singular values of regutm;(c) The relative errors of RRLSQR and LSQR

n	method	k_δ	Res	Err
200	LSQR	19	9.3359e-3	1.3424e-1
	RRLSQR	90	8.7893e-3	6.3682e-2
400	LSQR	35	8.5513e-3	1.1641e-1
	RRLSQR	134	9.1067e-3	5.7006e-2
800	LSQR	62	9.1109e-3	8.2321e-2
	RRLSQR	211	9.6155e-3	3.2082e-2
1000	LSQR	72	8.5872e-3	7.2140e-2
	RRLSQR	314	9.7773e-3	4.0888e-2

Table 1. The comparison between LSQR and RRLSQR

range restricted LSQR method, and consider the regularization properties of the RRLSQR method and present the regularized RRLSQR method.

In the analysis of the RRLSQR method, the $sin\Theta$ theorems for the 2-norm distances such as $\|sin\Theta(\mathcal{Q}_k, \mathcal{V}_k)\|$ and $\|sin\Theta(\mathcal{P}_k, \mathcal{U}_k)\|$ are established, where $\mathcal{Q}_k = K_k(A^TA, A^TAb)$ and $\mathcal{P}_k = K_k(AA^T, Ab)$ are the k dimensional right and left Krylov subspaces generated by Lanczos bidiagonalization, and $\mathcal{V}_k$ and $\mathcal{U}_k$ are the k dimensional dominant right and left singular subspaces of A. Furthermore, the accurate estimates on the distances for the three kinds of ill-posed problems are derived in the simple singular value case. These theorems show that Krylov subspace $\mathcal{Q}_k$ captures more information about $\mathcal{V}_k$ than Krylov subspace $\mathcal{V}_k^R = K_k(A^TA, A^Tb)$. To extend our results to the multiple singular value case will constitute our future work.

Acknowledgments

The research is supported by the National Natural Science Foundation of China under grant No.11571171. The authors are very grateful to the anonymous reviewer for his useful comments and suggestions which improved the representation of this paper.

References

1. M. Bellalij, L. Reichel, and H. Sadok. Some properties of range restricted GMRES methods. *Journal of Computational and Applied Mathematics* , 290:310–318, 2015.
2. Å. Björck. A bidiagonalization algorithm for solving large and sparse ill-posed systems of linear equations. *BIT Numerical Mathematics*, 28:659–670, 1988.
3. Å. Björck. *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia, 1996.
4. T. Buzug. *Computed Tomography: from Photon Statistics to Modern Cone-Beam CT*. Springer-Verlag, Berlin, 2008.

5. D. Calvetti, B. Lewis, and L. Reichel. GMRES-type methods for inconsistent systems. *Linear Algebra and Its Applications*, 316:157–169, 2000.
6. D. Calvetti, B. Lewis, and L. Reichel. On the choice of subspace for iterative methods for linear discrete ill-posed problems. *International Journal of Applied Mathematics and Computer Science*, 11:1069–1092, 2001.
7. D. Calvetti, B. Lewis, and L. Reichel. GMRES, L-curve, and discrete ill-posed problems. *BIT Numerical Mathematics*, 42:44–65, 2002.
8. D. Calvetti, B. Lewis, and L. Reichel. On the regularizing properties of the GMRES method. *Numerische Mathematik*, 91:605–625, 2002.
9. D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari. Tikhonov regularization and the L-curve for large discrete ill-posed problems. *Journal of Computational and Applied Mathematics* , 123:423–446, 2000.
10. D. Calvetti and L. Reichel. Tikhonov regularization of large linear problems. *BIT Numerical Mathematics*, 43:263–283, 2003.
11. N. Cesa-Bianchi. Applications of regularized least squares to pattern classification. *Theoretical Computer Science*, 382:221–231, 2007.
12. J. Chung, J. G. Nagy, and D. P. O'Leary. A weighted GCV method for lanczos hybrid regularization. *Electronic Transactions on Numerical Analysis* , 28:149–167, 2008.
13. J. Chung and K. Palmer. A hybrid LSMR algorithm for large-scale Tikhonov regularization. *SIAM Journal on Scientific Computing*, 37:562–580, 2015.
14. L. Dou and R. J. W. Hodgson. Application of the regularization methods to the inverse black body radiation problem. *IEEE Transactions on Antennas and Propagation*, 40:1249–1253, 1992.
15. L. Dykes, F. Marcellan, and L. Reichel. The structure of iterative methods for symmetric linear discrete ill-posed problem. *BIT Numerical Mathematics*, 54:129–145, 2014.
16. S. Gazzola, S. Noschese, P. Novati, and L. Reichel. Arnoldi decomposition, GMRES, and preconditioning for linear discrete ill-posed problems. *Applied Numerical Mathematics*, 142:102–121, 2019.
17. S. Gazzola and P. Novati. Inheritance of the discrete Picard condition in Krylov subspace methods. *BIT Numerical Mathematics*, 56:893–918, 2016.
18. S. F. Gilyazov. Regularizing algorithms based on the conjugate gradient method. *USSR Computational Mathematics and Mathematical Physics*, 26:8–13, 1986.
19. G. H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21:215–223, 1979.
20. M. Hanke. *Conjugate Gradient Type Methods for Ill-Posed Problems*. Longman, Essex, 1995.
21. M. Hanke and P. C. Hansen. Regularization methods for large-scale problems. *Surveys on Mathematics for Industry*, 3:253–315, 1993.
22. P. C. Hansen. The truncated SVD as a method for regularization. *BIT Numerical Mathematics*, 17:534–553, 1987.
23. P. C. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34:561–580, 1992.
24. P. C. Hansen. Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems. *Numerical Algorithms*, 6:1–35, 1994.
25. P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems*. SIAM, Philadelphia, 1998.
26. P. C. Hansen. *Discrete Inverse Problems: Insight and Algorithms*. SIAM, Philadelphia, 2010.
27. P. C. Hansen, Y. Dong, and K. Abe. Hybrid enriched bidiagonalization for discrete ill-posed problems. *Numerical Linear Algebra with Applications*, 26:e2230, 2019.
28. P. C. Hansen and T. K. Jensen. Noise propagation in regularizing iterations for image deblurring. *Electronic Transactions on Numerical Analysis* , 31:204–220, 2008.
29. B. Hofmann. *Regularization for Applied Inverse and Ill-posed Problems*. B. G. Teubner Verlag, Leipzig, 1986.
30. Y. Huang and Z. Jia. On regularizing effects of MINRES and MR-II for large scale symmetric discrete ill-posed problems. *Journal of Computational and Applied Mathematics* , 320:145–163, 2017.
31. Y. Huang and Z. Jia. Some results on the regularization of LSQR for large-scale discrete ill-posed problems. *Science China Mathematics*, 60:701–718, 2017.
32. T. K. Jensen and P. C. Hansen. Iterative regularization with minimum-residual methods. *BIT Numerical Mathematics*, 47:103–120, 2007.
33. Z. Jia. Approximation accuracy of the krylov subspaces for linear discrete ill-posed problems. *Journal of Computational and Applied Mathematics* , 374:112786, 2020.
34. Z. Jia. The low rank approximations and ritz values in lsqr for linear discrete ill-posed problems. *Inverse Problems*, 36:045013, 2020.
35. M. E. Kilmer, P. C. Hansen, and M. Español. A projection-based approach to general form tikhonov regularization. *SIAM Journal on Scientific Computing*, 29:315–330, 2007.
36. M. E. Kilmer and D. P. O'Leary. Choosing regularization parameters in iterative methods for ill-posed problems. *SIAM Journal on Matrix Analysis and Applications*, 22:1204–1221, 2001.
37. M. E. Kilmer and G. W. Stewart. Iterative regularization and minres. *SIAM Journal on Matrix Analysis and Applications*, 21:613–628, 1999.
38. A. Kirsch. *An Introduction to the Mathematical Theory of Inverse problems*. Springer, New York, 2011.
39. P. K. Kythe and P. Puri. *Computational Methods for Linear Integral Equations*. Birkhäuser, Boston, 2002.
40. C. L. Lawson and R. J. Hanson. *Solving Least Squares Problems*. Prentice-Hall, Englewood Cliffs, 1974.
41. L. Ma, T. Zeng, and G. Li. Hybrid variational model for texture image restoration. *East Asian Journal on Applied Mathematics*, 7:629–642, 2017.
42. V. A. Morozov. *Methods for Solving Incorrectly Posed Problems*. Springer, New York, 1984.
43. F. Natterer. *The Mathematics of Computerized Tomography*. SIAM, Philadelphia, 2001.
44. A. Neuman, L. Reichel, and H. Sadok. Algorithms for range restricted iterative methods for linear discrete ill-posed problems. *Numerical Algorithms*, 59:325–331, 2012.

45. A. Neuman, L. Reichel, and H. Sadok. Implementations of range restricted iterative methods for linear discrete ill-posed problems. *Linear Algebra and Its Applications*, 436:3974–3990, 2012.
46. D. P. O'Leary and J. A. Simmons. A bidiagonalization-regularization procedure for large scale discretizations of ill-posed problems. *SIAM Journal on Scientific Computing* , 2:474–489, 1981.
47. C. C. Paige and M. A. Sanders. Solution of sparse indefinite systems of linear equations. *SIAM Journal on Numerical Analysis*, 12:617–629, 1975.
48. C. C. Paige and M. A. Sanders. LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software*, 8:43–71, 1982.
49. R. Plato. Optimal algorithms for linear ill-posed problems yield regularization methods. *Numerical Functional Analysis and Optimization* , 11:111–118, 1990.
50. Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific Computing*, 7:856–869, 1986.
51. G. W. Stewart. *Matrix Algorithm I: Basic Decompositions*. SIAM, Philadelphia, 1998.
52. G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*. Academic Press, Boston, 1990.
53. A. N. Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet Mathematics - Doklady*, 4:1035–1038, 1963.
54. A. van der Sluis and H. A. van der Vorst. The rate of convergence of conjugate gradients. *Numerische Mathematik* , 48:543–560, 1986.
55. X. Zhang, L. Cheng, D. Chu, L.-Z. Liao, M. K. Ng, and R. C. E. Tan. Incremental regularized least squares for dimensionality reduction of large-scale data. *SIAM Journal on Scientific Computing*, 38:B414–B439, 2016.

