Effective Numerical Schemes for Ill-Posed Inverse Problems

Gundeep Singh, Jannatul Chhoa & Andreas Mang
Department of Mathematics, University of Houston, Houston, TX, USA

UNIVERSITY of HOUSTON

Teaser: Our goal is the design and analysis of effective numerical schemes for solving linear inverse problems of the form $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}_{\text{obs}}$. We extend our prior work presented in [1].

Mathematical Problem Formulation

We assume that the observed data $\boldsymbol{y}_{\text{obs}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\eta}$, where $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m,n}$, $\boldsymbol{y}_{\text{obs}} \in \mathbb{R}^m$, and $\boldsymbol{\eta} \propto \mathcal{N}(0, \boldsymbol{I}_n)$ is a random perturbation. In the inverse problem, we seek $\boldsymbol{x}$ given $\boldsymbol{y}_{\text{obs}}$ and $\boldsymbol{A}$ [2]. In general, $\boldsymbol{A}$ will not be invertible and $\boldsymbol{y}_{\text{obs}} \notin \operatorname{col} \boldsymbol{A}$. We can formulate the solution of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}_{\text{obs}}$ as a **regularized least squares problem** (**RLSQ**) of the form

$$\underset{\boldsymbol{x} \in \mathbb{R}^n}{\text{minimize}}\; f(\boldsymbol{x}), \quad \text{where } f(\boldsymbol{x}) \coloneqq \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}_{\text{obs}}\|_2^2 + \frac{\alpha}{2}\|\boldsymbol{L}\boldsymbol{x}\|_2^2. \quad (1)$$

The first term of f measures the discrepancy between the model prediction $\boldsymbol{y}_{\text{pred}} \coloneqq \boldsymbol{A}\boldsymbol{x}$ and $\boldsymbol{y}_{\text{obs}}$. The second term is a (Tikhonov-type) **regularization functional** with regularization operator $\boldsymbol{L} \in \mathbb{R}^{n,n}$ and regularization parameter $\alpha > 0$. This regularization model is introduced to alleviate mathematical issues with the **ill-posedness** of the inverse problem [3, 2]. We will see that the choices for $\boldsymbol{L}$ and α greatly affect the computed solution $\boldsymbol{x}_{\text{sol}}$ of (1). We consider the following regularization operators: (i) $\boldsymbol{L}^\mathsf{T}\boldsymbol{L} = \boldsymbol{I}_n$ (identity) and (ii) $\boldsymbol{L}^\mathsf{T}\boldsymbol{L} = -\Delta$ (Laplace operator).

Numerical Methods

Optimality Conditions

For an admissible solution $\boldsymbol{x}_{\text{sol}} \in \mathbb{R}^n$ of (1) we require that the first derivative $\nabla f(\boldsymbol{x}_{\text{sol}})$ of f vanishes, i.e., $\nabla f(\boldsymbol{x}_{\text{sol}}) = 0$. We have $\nabla f(\boldsymbol{x}) = \boldsymbol{A}^\mathsf{T}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}_{\text{obs}}) + \alpha \boldsymbol{L}^\mathsf{T}\boldsymbol{L}\boldsymbol{x}$. The first-order optimality conditions are

$$\boldsymbol{A}^\mathsf{T}(\boldsymbol{A}\boldsymbol{x}_{\text{sol}} - \boldsymbol{y}_{\text{obs}}) + \alpha \boldsymbol{L}^\mathsf{T}\boldsymbol{L}\boldsymbol{x}_{\text{sol}} \overset{!}{=} 0. \quad (2)$$

This equation is referred to as the normal equation.

RLSQ & TSVD

We consider different approaches to solve $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}_{\text{obs}}$ for $\boldsymbol{x}$. We can directly solve the optimality system (2); the numerical solution is given by $\boldsymbol{x}_{\text{sol}} = (\boldsymbol{A}^\mathsf{T}\boldsymbol{A} + \alpha \boldsymbol{L}^\mathsf{T}\boldsymbol{L})^{-1}\boldsymbol{A}^\mathsf{T}\boldsymbol{y}_{\text{obs}}$. Alternatively, we can compute the pseudo-inverse $\boldsymbol{A}^+$ of $\boldsymbol{A}$ based on a **truncated singular value decomposition** (**TSVD**). That is, we compute the factorization $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\mathsf{T}$ of $\boldsymbol{A}$, where $\boldsymbol{U} \in \mathbb{R}^{m,m}$ and $\boldsymbol{V} \in \mathbb{R}^{n,n}$ are orthogonal matrices for the left- and right-singular vectors, respectively, and $\boldsymbol{S} \in \mathbb{R}^{m,n}$ is a diagonal matrix for the singular values.

Now suppose that $\boldsymbol{A}$ has rank $r \ll \min\{n, m\}$. Under this assumption, $\boldsymbol{U} = [\boldsymbol{u}_1 \dots \boldsymbol{u}_r\;\; \boldsymbol{u}_{r+1} \dots \boldsymbol{u}_m] \in \mathbb{R}^{m,m}$, $\boldsymbol{S} = \operatorname{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0) \in \mathbb{R}^{m,n}$, and $\boldsymbol{V} = [\boldsymbol{v}_1 \dots \boldsymbol{v}_r\;\; \boldsymbol{v}_{r+1} \dots \boldsymbol{v}_n] \in \mathbb{R}^{n,n}$. Consequently, we can decompose $\boldsymbol{A}$ into

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\mathsf{T} = \boldsymbol{U}_r\boldsymbol{S}_r\boldsymbol{V}_r^\mathsf{T} = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\mathsf{T},$$

where $\boldsymbol{U}_r = [\boldsymbol{u}_1 \dots \boldsymbol{u}_r] \in \mathbb{R}^{n,r}$, $\boldsymbol{V}_r = [\boldsymbol{v}_1 \dots \boldsymbol{v}_r] \in \mathbb{R}^{n,r}$, $\boldsymbol{S}_r = \operatorname{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{r,r}$. The pseudo-inverse is given by $\boldsymbol{A}^+ = \boldsymbol{V}_r\boldsymbol{S}_r^{-1}\boldsymbol{U}_r^\mathsf{T}$. It follows that the solution $\boldsymbol{x}_{\text{sol}}$ of our problem is given by

$$\boldsymbol{x}_{\text{sol}} = \boldsymbol{A}^+\boldsymbol{y}_{\text{obs}} = \boldsymbol{V}_r\boldsymbol{S}_r^{-1}\boldsymbol{U}_r^\mathsf{T}\boldsymbol{y}_{\text{obs}} = \sum_{i=1}^{r} \sigma_i^{-1}(\boldsymbol{u}_i^\mathsf{T}\boldsymbol{y}_{\text{obs}})\boldsymbol{v}_i. \quad (3)$$

If we do not know the rank r of $\boldsymbol{A}$ or the singular values do not decay completely to zero, we can compute a rank r approximation to $\boldsymbol{A}$, i.e., $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\mathsf{T} \approx \boldsymbol{U}_r\boldsymbol{S}_r\boldsymbol{V}_r^\mathsf{T}$. We illustrate this in **Figure 1**.

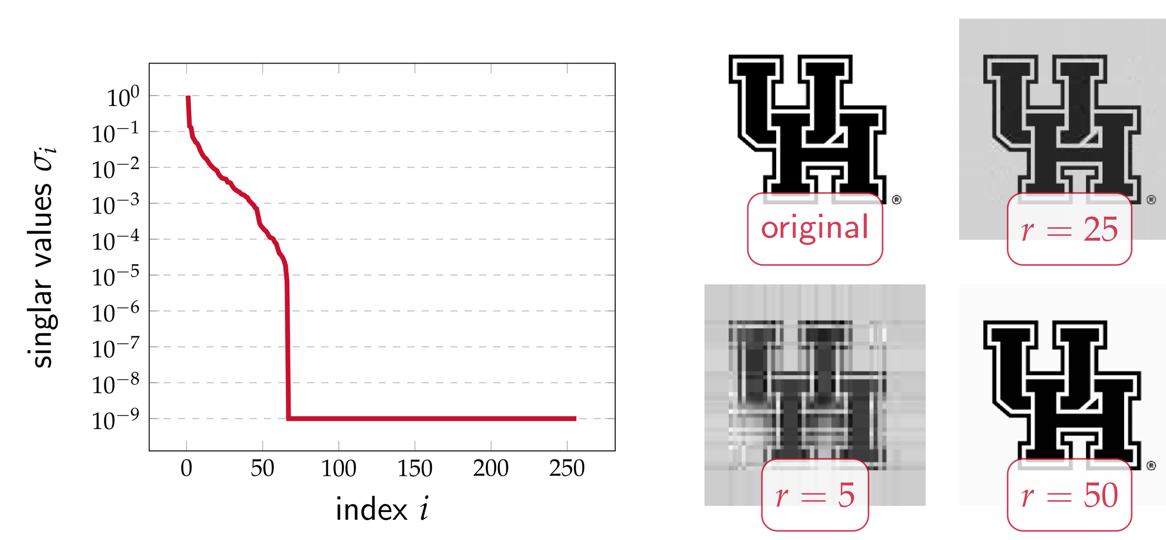

Figure 1: *Visualization of the compression of a matrix using low-rank approximations. We show the decay of the singular values for a considered matrix $\boldsymbol{A}$ of size 256×256 (UH logo; right: top left) on the left. The remaining figures (left: from top right to bottom right) show the reconstruction of the original matrix $\boldsymbol{A}$ using low rank approximations $\boldsymbol{U}_r\boldsymbol{S}_r\boldsymbol{V}_r^\mathsf{T}$ for different ranks $r \in \{5, 25, 50\}$.*

We can establish a connection between the TSVD and the RLSQ in (1) by introducing the regularized solution operator $\boldsymbol{R}_\alpha$,

$$\boldsymbol{R}_\alpha \boldsymbol{y}_{\text{obs}} = \sum_{i=1}^{n} g_\alpha(\sigma_i)(\boldsymbol{u}_i^\mathsf{T}\boldsymbol{y}_{\text{obs}})\boldsymbol{v}_i.$$

Here, g_α is a filter function. For the TSVD $g_\alpha(z) = 1/z$ for $z \geq \alpha$ and $g_\alpha(z) = 0$, otherwise. For the Tikhonov regularization operator $\boldsymbol{L}^\mathsf{T}\boldsymbol{L} = \boldsymbol{I}_n$, we obtain $g_\alpha(z) = z/(z^2 + \alpha)$, and, therefore,

$$\boldsymbol{x}_{\text{sol}} = \sum_{i=1}^{n} \frac{\sigma_i}{\sigma_i^2 + \alpha}(\boldsymbol{u}_i^\mathsf{T}\boldsymbol{y}_{\text{obs}})\boldsymbol{v}_i.$$

Randomized SVD

Computing the SVD for large-scale inverse problems can become computationally prohibitive. One way to alleviate the computational costs of constructing the low-rank approximation to $\boldsymbol{A}$ is to consider randomized algorithms [4, 5]. We are going to consider a prototype implementation of a randomized SVD. The pseudocode for this algorithm is given below.

```
procedure rSVD(A, r)
    draw random matrix Ω ∈ R^{n,r}
    Y ← AΩ
    Q ← qr_econ(Y)
    B ← Q^T A
    [Ũ, S, V] ← svd_econ(B)
    U ← QŨ
```

Numerical Experiments

In this section we present numerical experiments for the procedure detailed above. We compare the TSVD to the regularized least squares solution.

Synthetic Test Problem

We consider a synthetic test problem to study the performance of the proposed methodology. The operators in (1) are as follows: For $\boldsymbol{A}$, we consider a Helmholtz-type operator of the general form $\boldsymbol{A} = (-\Delta + k^2\boldsymbol{I}_n)^{-1}$, where $-\Delta$ is a Laplace operator, $k > 0$, and $\boldsymbol{I}_n$ is an $n \times n$ identity matrix. We compute $\boldsymbol{y}_{\text{obs}}$ by applying the forward operator $\boldsymbol{A}$ to $\boldsymbol{x}_{\text{true}}$. That is, $\boldsymbol{y}_{\text{obs}} = \boldsymbol{A}\boldsymbol{x}_{\text{true}} + \kappa\boldsymbol{\eta}$, $\boldsymbol{\eta} \propto \mathcal{N}(0, \boldsymbol{I}_n)$, $\kappa = \theta\|\boldsymbol{x}_{\text{true}}\|_2^2$, $\theta \in [0, 1]$. We select

$$\boldsymbol{x}_{\text{true}} \coloneqq (\sin(\boldsymbol{z}) + \gamma\boldsymbol{z} \odot \sin(4\boldsymbol{z})) \odot \exp(-\|\boldsymbol{z} - \pi\|_2^2/2\kappa),$$

with $\kappa = 9/10$ and $\gamma = 7/2$ and $z_i = hi$, $h = 2\pi/n$, $i = 1, \dots, n$.

Numerical Results

We report numerical results for different strategies to solve the considered inverse problem for the test problem described in the former section in **Figure 2**.

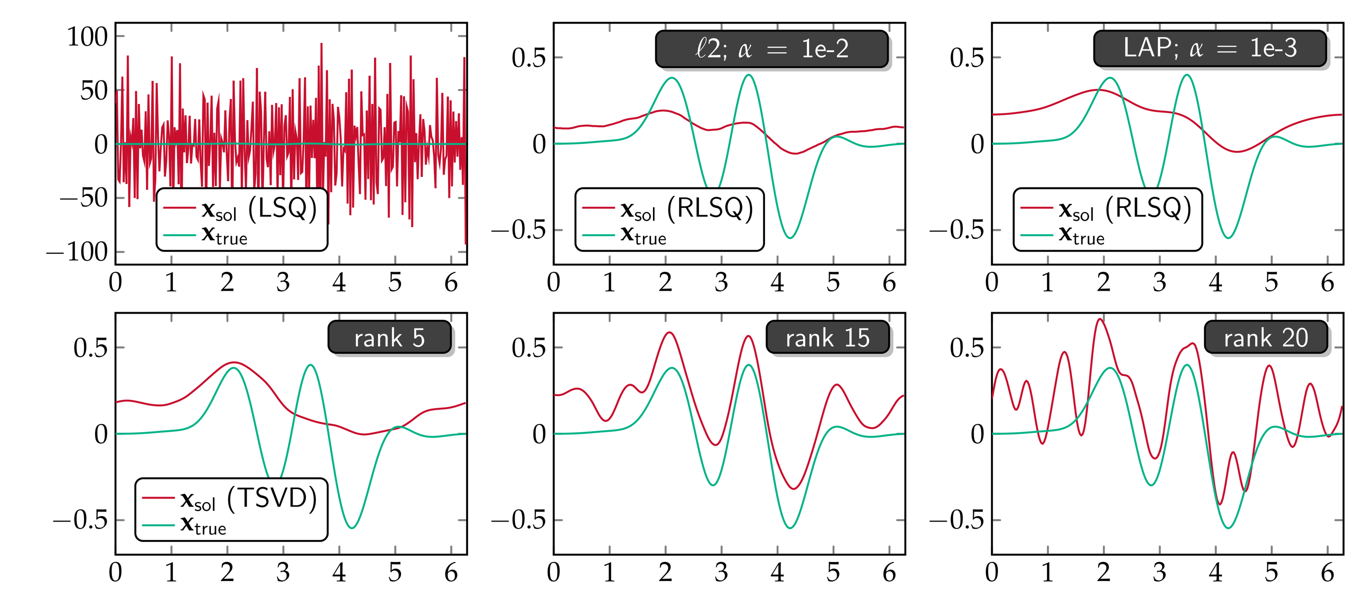

Figure 2: *We report solutions for the RLSQ in (1) for different regularization operators and compare them to the results obtained for the TSVD. The numerical solution is shown in red and the true solution $\boldsymbol{x}_{true}$ is shown in green. Top row: For the unregularized case ($\alpha = 0$) we can observe that the noise is amplified; the computed solution has nothing to do with the true solution. For the regularized case (with regularization parameters α = 1e-2/1e-3 for the regularization operators $\boldsymbol{L}^\mathsf{T}\boldsymbol{L} = \boldsymbol{I}_n / -\Delta$, we can see that we underfit the data. We show results for the TSVD in the bottom row for different target ranks r. We consider a randomized SVD for computing these results. The best result is obtained by computing the solution through a low rank approximation $\boldsymbol{U}_r\boldsymbol{S}_r\boldsymbol{V}_r^\mathsf{T} \approx \boldsymbol{A}$ for a target rank $r = 15$.*

Conclusions

We have developed and tested a computational framework for solving and regularizing linear inverse problems [2]. We have compared results for different variants of Tikhonov-type regularization operators to those obtained by a TSVD. To construct the TSVD we have considered efficient randomized algorithms [4, 5].

References

[1] H. Rosso & A Mang. *Regularization Schemes for Linear Inverse Problems*, Poster Presentation, Undergraduate Research Day, University of Houston, 2021.

[2] C. Vogel. *Computational Methods for Inverse Problems*, SIAM, 2002.

[3] H. W. Engl, M. Hanke & A. Neubauer, *Regularization of Inverse Problems*, Springer, 2000.

[4] N. Halko, P.-G. Martinsson & J. A. Tropp. *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*. SIAM Review 53(2):217–288, 2011.

[5] P. G. Martinsson & J. A. Tropp. *Randomized numerical linear algebra: Foundations and algorithms*. Acta Numerica 29, 403–572, 2020.