Improving LSQR with oversampling: application for inverse problems

Rosemary Renaut[1] Anthony Helmstetter[1] Saeed Vatankhah[2]

1: School of Mathematical and Statistical Sciences, Arizona State University, renaut@asu.edu,

2: Institute of Geophysics, University of Tehran,
Hubei Subsurface Multi-scale Imaging Key Laboratory,
Institute of Geophysics and Geomatics,
China University of Geosciences, Wuhan, China

Modern Challenges in Imaging: In the Footsteps of Allan MacLeod Cormack

Outline

Motivation and Background

- Inversion Undersampled Magnetic / Gravity Data
- Basic technique: the singular value decomposition (SVD)
- Focusing Inversion: Reweighted Regularization [LK83, WR07]

Numerical Methods for Large Scale: Approximating the SVD

- Krylov: Golub Kahan Bidiagonalization - LSQR [PS82]
- Randomized SVD [HMT11]
- Enhancing LSQR by Oversampling: SVDS [Lar98, BR05]

Properties and Simulations

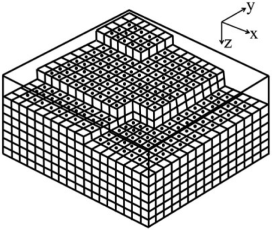
- Contrast Hybrid SVDS
- Angles between singular vectors
- Angles between subspaces
- Image Restoration with Focusing Inversion
- Undersampled Focusing Inversion of Geophysical Data

Conclusions and Future Work

Motivation Example: Large Scale $3D$ Magnetic Anomaly Inversion

Observation point $\mathbf{r} = (x, y, z)$

Vertical magnetic anomaly $m(\mathbf{r})$ is given using Biot-Savart Law



$$m(\mathbf{r}) \quad \propto \quad \int_{d\Omega} K(\mathbf{r}, \mathbf{r}')\kappa(\mathbf{r}')d\Omega$$

Susceptibility $\kappa(\mathbf{r}')$ at $\mathbf{r}' = (x', y', z')$

Linear Relation $\boxed{\mathbf{m} = G\boldsymbol{\kappa}}$ (or $\mathbf{b} = A\mathbf{x}$)

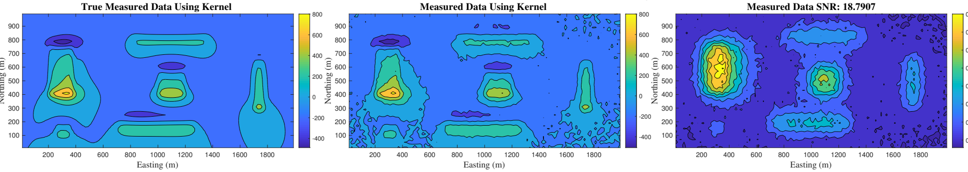
Aim: Given surface observations m_{ij} find susceptibility κ_{ijk}

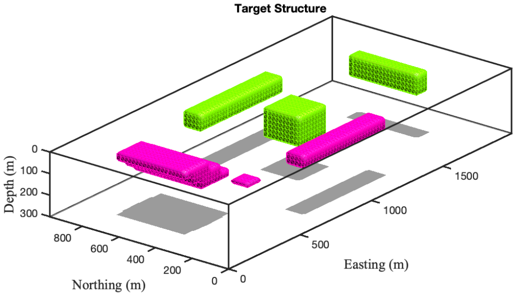
Underdetermined, measurements 5000, unknowns 75000

Practical Approaches for Large Scale Ill-Posed Problems needed

Magnetic and Gravity data $m = 5000$, $n = 75000$ SNR: 19 [VRA18]

The Model, True Data and Noisy Data





Example Results: Magnetic for subspace size k

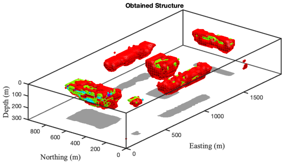


Figure: LSQR $k = 15$, 1s

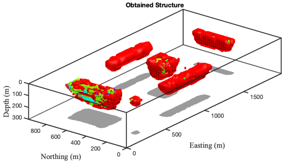


Figure: LSQROS 10% $k = 15$, 1s

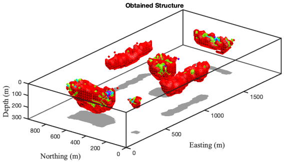


Figure: RSVD $k = 1000$, 10% 1073s

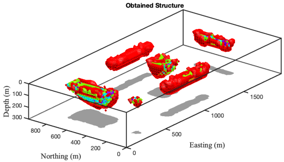


Figure: RSVD: power iteration $k = 1000$, 10% 2211s

Notation: Spectral Decomposition of the Solution: The SVD

Consider general discrete problem

$$A\mathbf{x} = \mathbf{b}, \quad A \in \mathbb{R}^{m \times n}, \quad \mathbf{b} \in \mathbb{R}^m, \quad \mathbf{x} \in \mathbb{R}^n.$$

Singular value decomposition (SVD) of A rank $r \leq \min(m, n)$

$$A = U\Sigma V^T = \sum_{i=1}^{r} \mathbf{u}_i \sigma_i \mathbf{v}_i^T, \quad \Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_r).$$

Singular values σ_i, singular vectors $\mathbf{u}_i$, $\mathbf{v}_i$, rank r.
Expansion for the solution:

$$\mathbf{x} = \sum_{i=1}^{r} \frac{\mathbf{s}_i}{\sigma_i} \mathbf{v}_i, \quad \mathbf{s}_i = \mathbf{u}_i^T \mathbf{b}$$

Regularization: Filtering and Truncation

Filtered and Truncated solution

$$\mathbf{x} = \sum_{i=1}^{k} \gamma_i(\alpha) \frac{\mathbf{s}_i}{\sigma_i} \mathbf{v}_i, \quad \gamma_i(\alpha) = \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2}, \; i = 1 \ldots k,$$

Solves Standard Form

$$\begin{aligned} \mathbf{x}(\alpha) &= \underset{\mathbf{x}}{\operatorname{argmin}}\{\|\mathbf{b} - A\mathbf{x}\|^2 + \alpha^2\|\mathbf{x}\|^2\} \\ \mathbf{x}_k(\alpha) &\approx \underset{\mathbf{x}}{\operatorname{argmin}}\{\|\mathbf{b} - A_k\mathbf{x}\|^2 + {\alpha_k}^2\|\mathbf{x}\|^2\} \end{aligned}$$

Generalized Tikhonov - L invertible (transfer to standard form)

$$\begin{aligned} \mathbf{x}(\alpha) &= \underset{\mathbf{x}}{\operatorname{argmin}}\{\|\mathbf{b} - A\mathbf{x}\|^2 + \alpha^2\|L\mathbf{x}\|^2\} \\ \mathbf{x}_k(\alpha) &\approx L^{-1}(\underset{\mathbf{y}}{\operatorname{argmin}}\{\|\mathbf{b} - A_k L^{-1}\mathbf{y}\|^2 + {\alpha_k}^2\|\mathbf{y}\|^2\}) \end{aligned}$$

Iterative Reweighted Regularization: Focusing Inversion [LK83, WR07] with iteration count t:

$$\|A\mathbf{x} - \mathbf{b}\|^2 + \alpha^2 \|L^{(t)}(\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)})\|^2, \quad t > 0$$

Regularization operator $L^{(t)}$. ϵ ensures $L^{(t)}$ invertible

$$(L^{(t)})_{ii} = ((\mathbf{x}_i^{(t-1)} - \mathbf{x}_i^{(t-2)})^2 + \epsilon)^{-1/4} \quad \epsilon > 0$$

Invertibility use $(L^{(t)})^{-1}$ as right preconditioner for A

$$(L^{(t)})_{ii}^{-1} = ((\mathbf{x}_i^{(t-1)} - \mathbf{x}_i^{(t-2)})^2 + \epsilon)^{1/4} \quad \epsilon > 0$$

Regularization parameter α_k automatic update each t.

Cost of $L^{(t)}$ is minimal: it is diagonal

Generalized Tikhonov regularization: System $A_k (L^{(t)})^{-1}$

Iterative Krylov Method: LSQR Approximates SVD [PS82]

- Define $\beta_1 := \|\mathbf{b}\|_2$, $\mathbf{e}_1^{(k+1)}$ first column of I_{k+1} and $\beta_1 H_{k+1}\mathbf{e}_1^{(k+1)} = \mathbf{b}$
- Factorize $AG_k = H_{k+1}B_k$, Lanczos vectors G_k and H_{k+1}
- Lanczos vectors span $\mathcal{K}_{k+1}\{AA^T, \mathbf{b}\}$ and $\mathcal{K}_k\{A^TA, A^T\mathbf{b}\}$. and are column orthogonal. $B_k \in \mathcal{R}^{(k+1)\times k}$ is lower bidiagonal.
- Projected Problem

$$B_k\mathbf{w}_k \approx \beta_1\mathbf{e}_1^{(k+1)}, \quad \mathbf{x}_k = G_k\mathbf{w}_k$$

- Hybrid projected problem

$$\mathbf{x}_k = G_k\left(\operatorname{argmin}\{\|B_k\mathbf{w}_k - \beta_1\mathbf{e}_1^{(k+1)}\|^2 + \alpha^2\|\mathbf{w}_k\|^2\}\right)$$

- Solution defined by SVD of $B_k = \tilde{U}\tilde{\Sigma}\tilde{V}^T$
- Ritz vectors, columns of $G_k\tilde{V}$ and $H_{k+1}\tilde{U}$, give $\tilde{A}_k$.

Approximate SVD: $\boxed{\tilde{A}_k} = (H_{k+1}\tilde{U})\tilde{\Sigma}(G_k\tilde{V})^T$

Randomized Singular Value Decomposition : Proto [HMT11]

$A \in \mathcal{R}^{m\times n}$, target rank k, oversampling parameter p, $k+p \ll m$, $m \geq n$. Power factor q. Compute $A \approx \boxed{\overline{A}_k} = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T$.

1: Generate a Gaussian random matrix $\Omega \in \mathcal{R}^{n\times(k+p)}$.
2: Compute $Y = A\Omega \in \mathcal{R}^{m\times(k+p)}$. $Y = \mathrm{qr}(Y)$
3: If $q > 0$ repeat q times $\{\ [\mathbf{Y}, \sim] = \mathrm{qr}(A^T \mathrm{qr}(A\mathbf{Y}))\}$ Power
4: Form $B = Y^T A \in \mathcal{R}^{(k+p)\times n}$. $(Q = Y)$
5: Find SVD $B = U_B \Sigma_B V_B^T$, $U_B \in \mathcal{R}^{(k+p)\times(k+p)}$, $V_B \in \mathcal{R}^{k\times k}$
6: $\overline{U}_k = QU_B(:, 1:k)$, $\overline{V}_k = V_B(:, 1:k)$, $\overline{\Sigma}_k = \Sigma_B(1:k, 1:k)$

- Hybrid projected problem

$$\mathbf{x}_k = \mathrm{argmin}\{\|\overline{A}_k \mathbf{x}_k - \mathbf{b}\|^2 + \alpha^2 \|\mathbf{x}_k\|^2\}$$

- Solution defined by approximation $\overline{A}_k = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T$

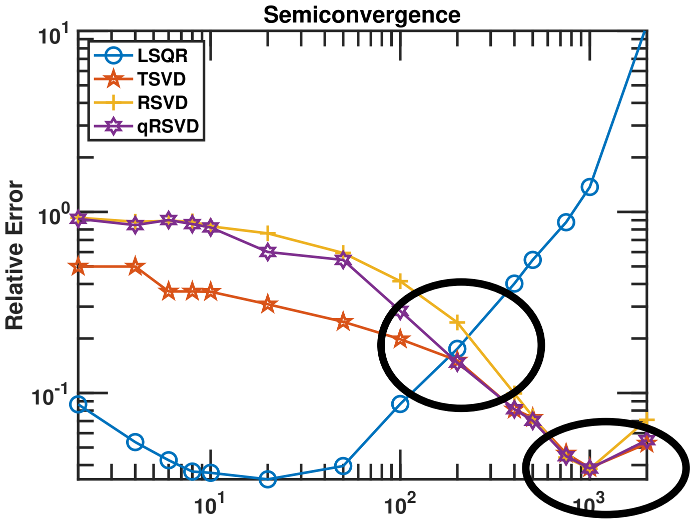
Approximate SVD $\boxed{\overline{A}_k} = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T$

Compute $\overline{U}_k \in \mathcal{R}^{m\times k}$, $\overline{\Sigma}_k \in \mathcal{R}^{k\times k}$, $\overline{V}_k \in \mathcal{R}^{n\times k}$.

1: Generate a Gaussian random matrix $\mathbf{\Omega} \in \mathcal{R}^{(k+p)\times m}$.
2: Compute matrix $\mathbf{Y} = \mathbf{\Omega} A \in \mathcal{R}^{(k+p)\times n}$.
3: Compute $\mathbf{Q} \in \mathcal{R}^{n\times(k+p)}$ via QR factorization $\mathbf{Y}^T = \mathbf{QR}$.
4: If $q > 0$ repeat q times $\{ [\mathbf{Q}, \sim] = \mathrm{qr}(A^T \mathrm{qr}(A\mathbf{Q}))\}$ Power
5: Form $B = A\mathbf{Q} \in \mathcal{R}^{m\times(k+p)}$ using factored form of $\mathbf{Q}$.
6: Compute the matrix $B^T B \in \mathcal{R}^{(k+p)\times(k+p)}$.
7: Compute the eigen-decomposition of $B^T B$;
$[\tilde{\overline{V}}_{k+p}, D_{k+p}] = \mathrm{eig}(B^T B)$.
8: Compute $\overline{V}_k = \mathbf{Q}\tilde{\overline{V}}_l(:, 1:k)$; $\overline{\Sigma}_k = \sqrt{D_l}(1:k, 1:k)$; and
$\overline{U}_k = B\tilde{\overline{V}}_k(:, 1:k)\overline{\Sigma}_k^{-1}$.

Yields $\boxed{\overline{A}_k} = \overline{U}_k \overline{\Sigma}_k \overline{V}_k^T$

Semi-convergence of LSQR, TSVD, RSVD and power RSVD $q = 1$

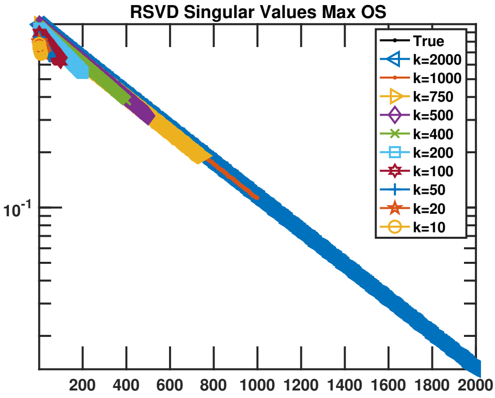


$\overline{A}_k$ inherits the ill-conditioning of A_k

AIM: optimal stable k

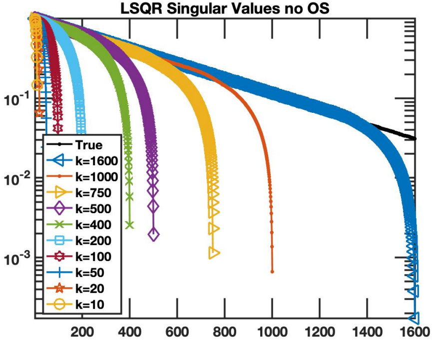
Theoretical Properties: Contrasting spectrum of Surrogates

Figure: RSVD: Good Approximation of Dominant Singular Values for a problem of size 4096×4096 using the RSVD algorithm using $\mathbf{100\%}$ oversampling, as compared to the exact singular values of the problem.



Theoretical Properties: Contrasting spectrum of Surrogates

Figure: LSQR: (with reorthogonalization) Good Approximation of fewer dominant singular values for a problem of size 4096×4096 using the LSQR algorithm with Krylov subspace of size k, as compared to the exact singular values of the problem.



The LSQR / RSVD spectrum: Key Properties

LSQR

- Good estimates of extremal singular values
- Interior eigenvalue approximations *improve* for increasing k
- Dominant spectrum **stabilizes**, increasing k.
- $\boxed{\tilde{A}_k}$ is not an approximation to $\boxed{A_k}$
- Ill-conditioning is captured.

RSVD

- Approximates dominant singular values with sufficient oversampling
- With power iteration improved $\boxed{\overline{A}_k} \approx \boxed{A_k}$
- Does not capture ill-conditioning.

LSQROS: Apply LSQR for Krylov space size $k+p$

- Use LSQR to space size $k+p$: $B_{k+p} \in \mathcal{R}^{(k+p+1)\times(k+p)}$.

$$AG_{k+p} = H_{k+p+1}B_{k+p}$$

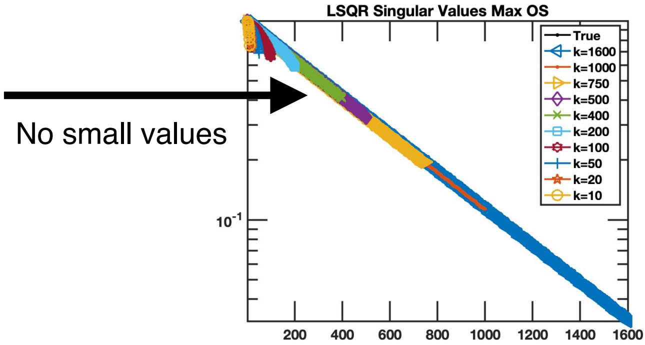
- Find SVD (Enlarged Krylov space)

$$\boxed{[\tilde{U}, \tilde{\Sigma}, \tilde{V}] = \mathrm{svd}(B_{k+p})}$$

- Truncate

$$\tilde{U} = \tilde{U}(:, 1:k), \quad \tilde{V} = \tilde{V}(:, 1:k), \quad \tilde{\Sigma} = \tilde{\Sigma}(1:k, 1:k)$$

Figure: LSQR add oversampling: Good Approximation of fewer dominant singular values for a problem of size 4096×4096 using the LSQR algorithm with extended Krylov subspace of size k, as compared to the exact singular values of the problem. Oversampled $\mathbf{100\%}$



Alternative view of oversampling: svds

- Apply svds find dominant singular space of size k:

$$\boxed{[\tilde{U}, \tilde{\Sigma}, \tilde{V}] = \text{svds}(B_{k+p}, k)}$$

  $\tilde{U} \in \mathcal{R}^{(k+p+1)\times k}$, $\tilde{V} \in \mathcal{R}^{(k+p)\times k}$, $\tilde{\Sigma}$ is size $k \times k$.
- Uses tolerances to determine required size of SVD from B_{k+p} that is needed.
- When p relatively large compared to k, the SVD may not be of size $k + p$.
- Use existing software: eg Propack: lansvd [Lar98, BR05].

Why not just abandon LSQR and apply svds directly to A?

For $A \in \mathcal{R}^{m\times n}$, target rank k, using space size $k+p$

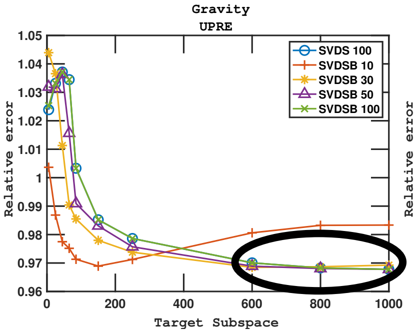
- Directly use svds to find dominant space of size k using extended Krylov space $k+p$.

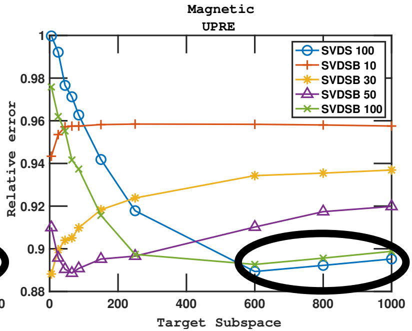
$$\boxed{[U_k, \Sigma_k, V_k] = \mathrm{svds}(A, k, '\mathrm{SubspaceDimension}', k+p)}$$

- Approximate SVD is immediate if tolerance is met.
- Can adjust p if tolerance is not met. e.g. iterate on $k+p$ to force acceptable tolerance for size k.

Why LSQROS and not $\mathrm{svds}(A)$?

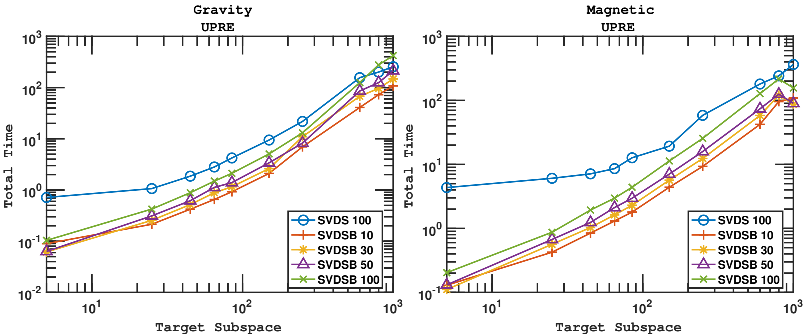
Hybrid SVDS: Problem Size 5000 by 75000 (One Step)





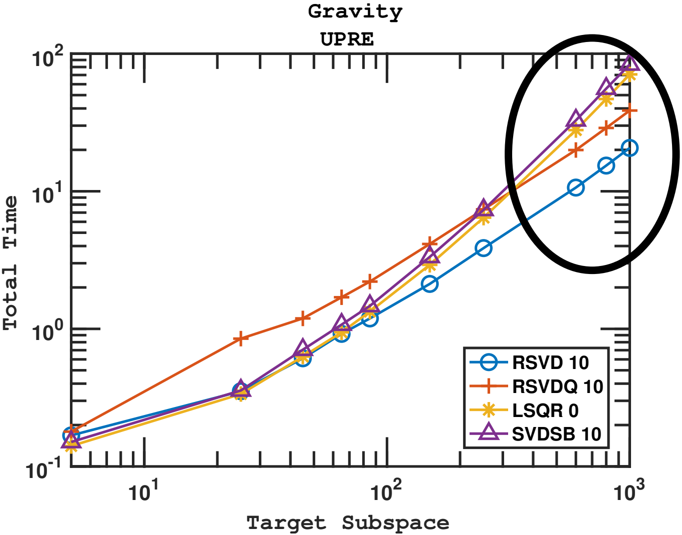
Comparison of relative error : SVDSB converges to SVDS

Hybrid SVDS: Problem Size 5000 by 75000 (One Step)



Comparison of time SVDS more expensive than SVDSB

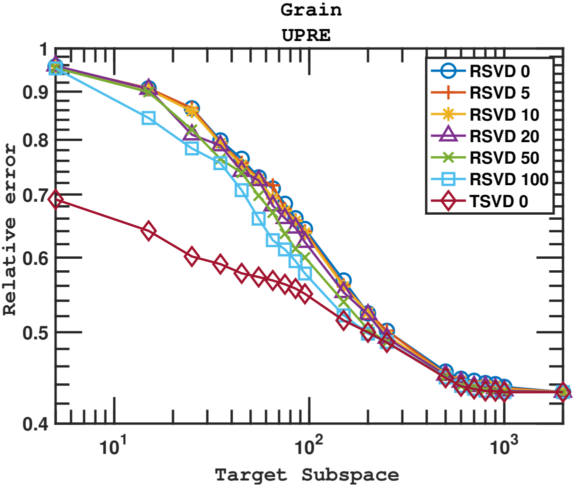
Gravity: Problem Size 5000 by 75000 (One Step)



Comparison of Time : SVDSB expensive increasing k

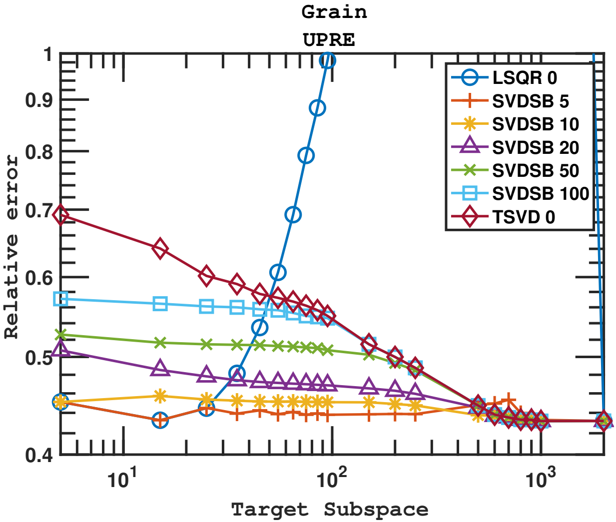
But LSQR also expensive as compared to RSVD

Hybrid RSVD: RSVD with regularization and oversampling



Relative Errors larger than TSVD for small k

Hybrid LSQR: regularization and oversampling image restoration



Relative Errors less than TSVD for small k

SVDSB reduces semiconvergence issue of LSQR

Preliminary Summary

SVDS / SVDSB

1. SVDS can be used in place of LSQR
2. SVDS can be applied directly to projected problem
3. SVDSB cheaper than SVDS.

RSVD / LSQR options

1. OS for LSQR is effective for small k
2. RSVD is not effective for small k

Spectrum approximation is not a sufficient guide for accuracy

Other Properties: rank k approximation of A

RSVD and LSQR provide approximate TSVD (see references)

	TSVD	LSQR	RSVD
Size	k	k	$k+p$
Model	$\boxed{A_k}$	$\boxed{\tilde{A}_k}$	$\boxed{\overline{A}_k}$
SVD	$U_k \Sigma_k V_k^T$	$(H_{k+1}\tilde{U})\tilde{\Sigma}(G_k\tilde{V})^T$	$\overline{U}_k \overline{\Sigma}_k \overline{V}_k{}^T$
Terms	k	k	k
$\mathbf{s}_i$	$\mathbf{u}_i^T\mathbf{b}$	$(H_{k+1}\tilde{U}_k)_i^T\mathbf{b}$	$\overline{\mathbf{u}}_i^T\mathbf{b}$
Basis	columns of V_k	columns of $G_k\tilde{V}_k$	columns of $\overline{V}_k$
$\Vert A - \boxed{A_k} \Vert$	σ_{k+1}	Theorem $\boxed{\tilde{A}_k}$	Theorem $\boxed{\overline{A}_k}$
$\sin(\langle V_k, \cdot \rangle)$	Golub [GvL96]	Delft et al [DDLT91]	Saibaba [Sai19]

Accuracy is well-studied

Theorem (RSVD Proto: with power iteration q [HMT11])

$$\mathrm{E}(\|A - \overline{A}_k\|) \leq \left(1 + \left[1 + 4\sqrt{\frac{2\min\{m,n\}}{k-1}}\right]^{1/(2q+1)}\right)\sigma_{k+1}$$

Theorem (LSQR [Jia17]: Fast decay of singular values $\sigma_i = \zeta\rho^{-i}$, $\rho > 2$ and noise level contaminates at coefficient ℓ)

$\tilde{A}_k = H_{k+1} B_k {G_k}^T$ *is a near best rank* k *approximation to* A *for* $k = 1, 2, \ldots, \ell - 1$.

Theorems on approximation of the spectral space: Angles between subspaces (vectors) formed by TSVD and approximate TSVD :

Theorem ([DDLT91]: For LSQR ($\sigma_i \neq \sigma_j$) and $\|A - \tilde{A}_k\| \leq \nu_k \|A\| = \nu_k \sigma_1$ If $2\nu_k < \min_{i \neq j} |\sigma_i - \sigma_j|$, then)

$$\mathit{max}(\sin\theta(\mathbf{u}_i, \tilde{\mathbf{u}}_i), \sin\theta(\mathbf{v}_i, \tilde{\mathbf{v}}_i)) \leq \frac{\nu_k}{\min_{i \neq j} |\sigma_i - \sigma_j| - \nu_k} \leq 1.$$

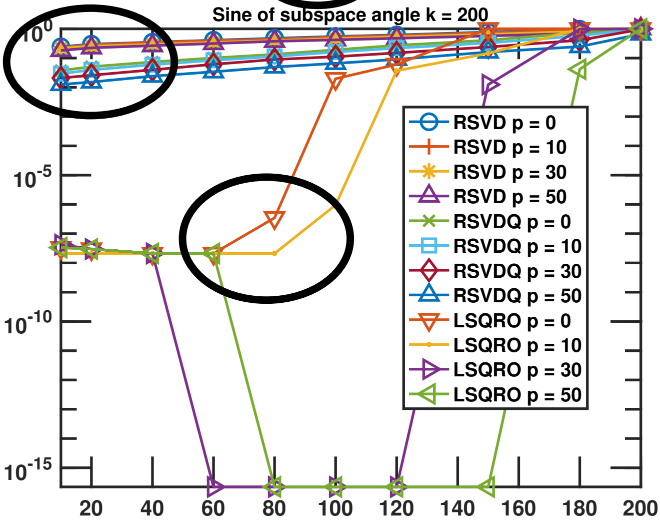
Theorem (Convergence of Lanczos Vectors [Saa11, Theorem 6.3]. $L_i(\sigma_n^2) = \prod_{j=1}^{i-1} \frac{(\sigma_j)^2 - \sigma_n^2}{(\sigma_j)^2 - \sigma_i^2}$. $\rho_i = \frac{\sigma_i^2 - \sigma_{i+1}^2}{\sigma_{i+1}^2 - \sigma_n^2}$, Chebyshev polynomial C_{k-i})

$$\tan(\Theta(\mathbf{v}_i, G_k)) \leq \frac{L_i(\sigma_n^2)}{C_{k-i}(1 + 2\rho_i)} \tan(\Theta(\mathbf{v}_i, G_1)).$$

Theorem ([Sai19, Theorem 4] RSVD canonical subspace angles $i = 1 : k$. $\gamma_i = \sigma_{k+1}/\sigma_i$)

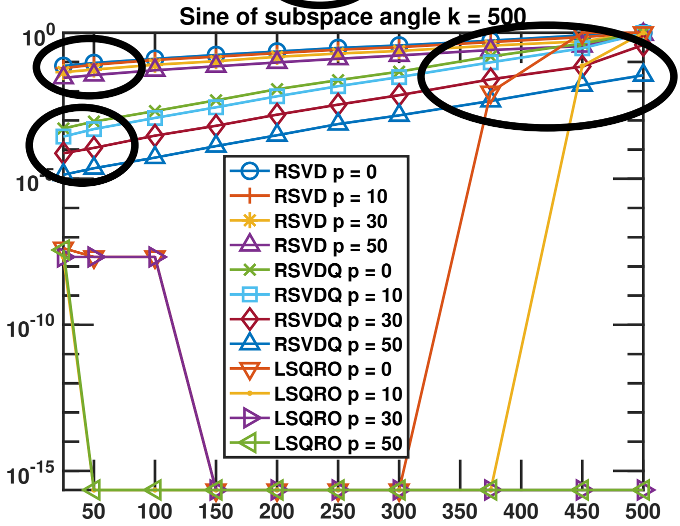
$$\mathit{max}(\sin\theta(\mathbf{u}_i, \overline{\mathbf{u}}_i), \sin\theta(\mathbf{v}_i, \overline{\mathbf{v}}_i)) \leq \gamma_i \frac{\gamma_k^{2q}}{1 - \gamma_k} \|(V_\perp^T \Omega)(V_k^T \Omega)^\dagger\|_2$$

Figure: Fixed $k = 200$ and increasing p



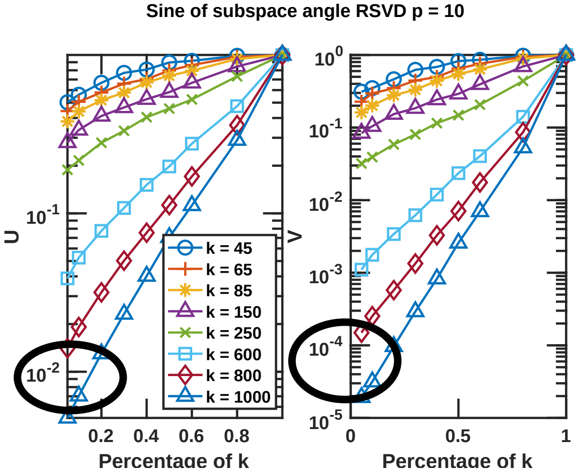
Contrasting sines of angles between singular vectors

Figure: Fixed $k = 500$ and increasing p



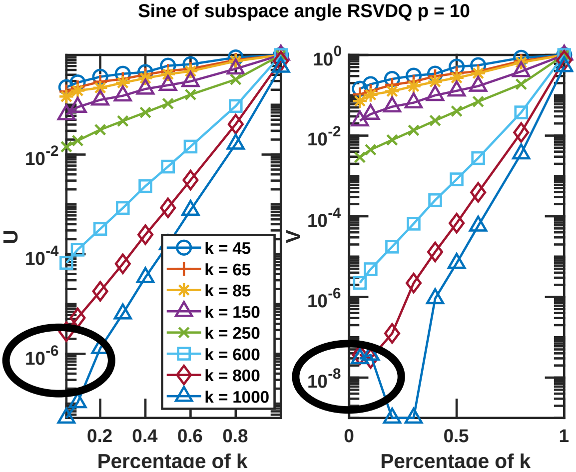
Contrasting Sines of Subspace Canonical Angles :

Figure: RSVD ($p = 10$): $\sin\Theta(U_\ell, \overline{U}_\ell)$, $\sin\Theta(V_\ell, \overline{V}_\ell)$ Increasing k: $\ell = fk$, f is percent



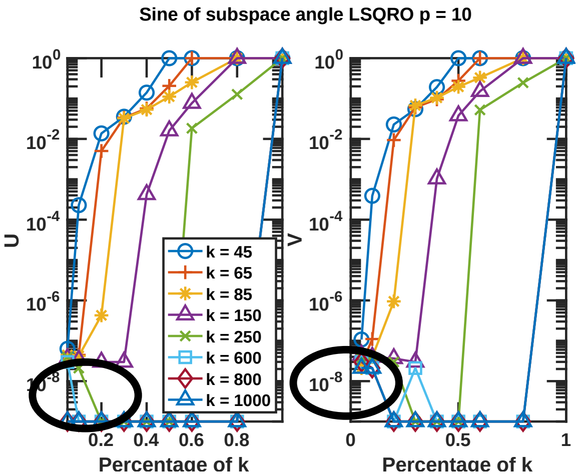
Contrasting Sines of Subspace Canonical Angles :

Figure: RSVDQ ($p = 10$): $\sin\Theta(U_\ell, \overline{U}_\ell)$, $\sin\Theta(V_\ell, \overline{V}_\ell)$ Increasing k: $\ell = fk$, f is percent



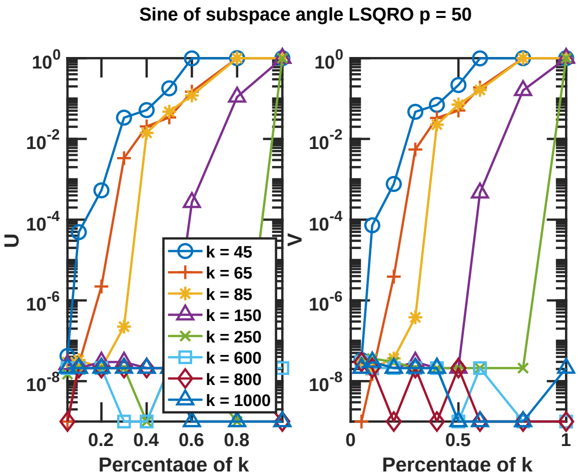
Contrasting Sines of Subspace Canonical Angles :

Figure: LSQROS ($p = 10$): $\sin\Theta(U_\ell, \tilde{U}_\ell)$, $\sin\Theta(V_\ell, \tilde{V}_\ell)$ Increasing k: $\ell = fk$, f is percent



Contrasting Sines of Subspace Canonical Angles :

Figure: LSQROS $(p = 50)$: $\sin\Theta(U_\ell, \tilde{U}_\ell)$, $\sin\Theta(V_\ell, \tilde{V}_\ell)$ Increasing k: $\ell = fk$, f is percent



Observations: LSQR and RSVD

1. Canonical angles between the singular vectors are far smaller for LSQR than RSVD and RSVDQ, particularly with oversampling.
2. Canonical angles between dominant subspaces are far smaller for LSQR than RSVD for equivalent small k
3. RSVD does not capture the subspace of rank k from a $k+p$ estimate as well as LSQROS - canonical angles are larger.
4. Subspace alignment stabilizes for LSQROS.
5. Conclude: LSQROS better mimics TSVD.

Restoration Grain size 256×256: SNR 20: Dominant space of size 500

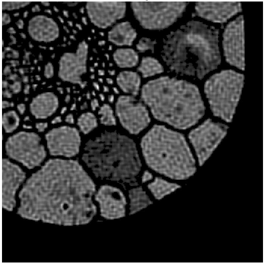


Figure: True Data

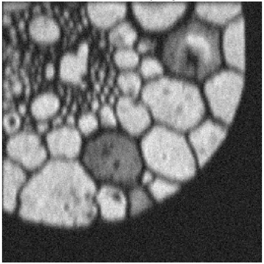


Figure: Blurred Noisy Data

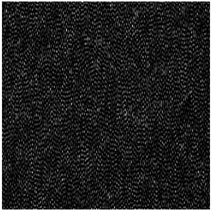


Figure: LSQR

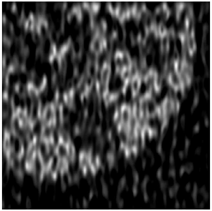


Figure: RSVDQ

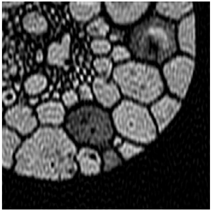
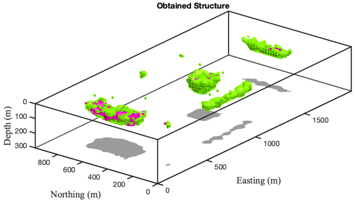


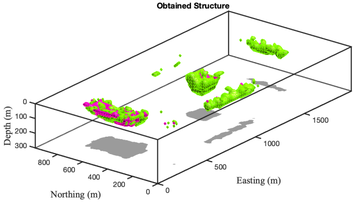
Figure: LSQROS

Stabilization with LSQROS and 10% oversampling

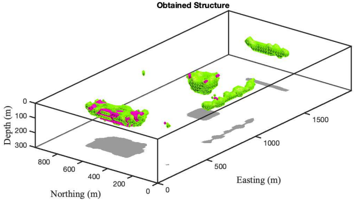
Gravity Results: Volume Rendering

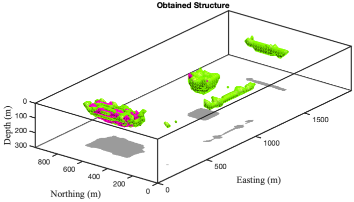
LSQR $k = 1000$ LSQROS, $k = 150$, 10% oversampling





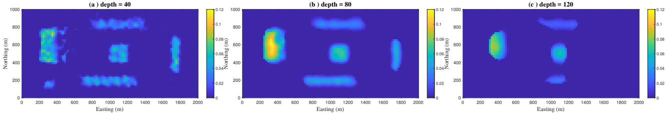
RSVD $k = 1000$ and RSVDQ $k = 1000$ 10% oversampling



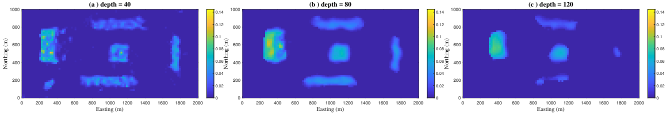


Gravity Results: UPRE for parameter estimation

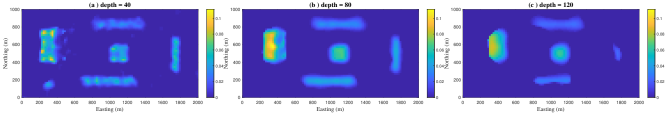
LSQR with $k = 1000$, time $1016s$, 0% oversampling (14 iterations)



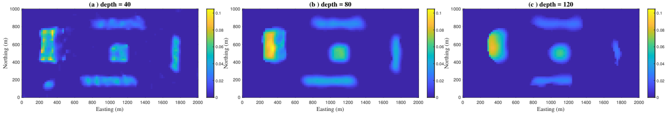
LSQROS with $k = 150$, time $43s$, 10% oversampling (14 iterations)



RSVD with $k = 1000$, time $339s$, 10% oversampling (15 iterations)

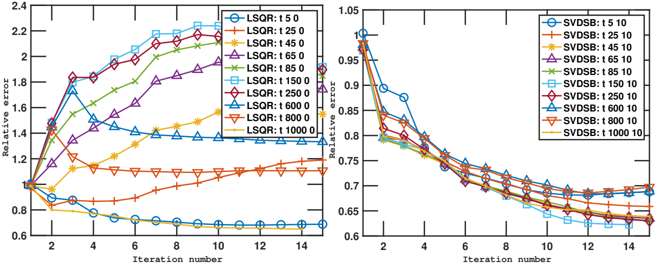


RSVDQ with $k = 1000$, time $537s$ 10% oversampling (13 iterations)

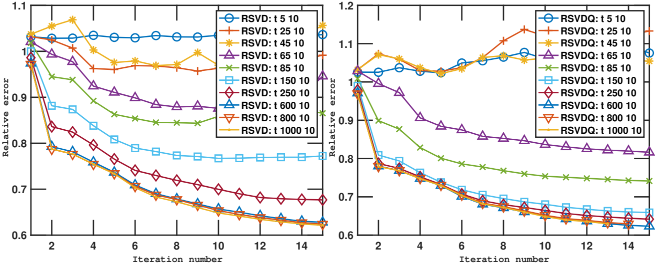


Gravity Results: Relative Error

LSQR and LSQROS 10% oversampling



RSVD and RSVDQ, 10% oversampling



Conclusions: LSQR with Extended Krylov

Canonical Angles Accurate dominate subspace is critical.

Extension of Krylov Space Improves dominant space accuracy.

RSVD / LSQR Trade offs depend on speed by which singular values decrease (degree of ill-posedness)

Cost While LSQROS more expensive than LSQR, provides the dominant subspace more accurately for p small.

Hybrid Implementations stabilize the solution errors.

Heuristics verified on a practical application.

Future

- Apply for Generalized Regularizers
- Stabilize RSVD oversampling choice using svds?

Some key references

James Baglama and Lothar Reichel.
Augmented implicitly restarted lanczos bidiagonalization methods.
SIAM Journal on Scientific Computing, 27(1):19–42, 2005.

Percy Deift, James Demmel, Luen-Chau Li, and Carlos Tomei.
The bidiagonal singular value decomposition and Hamiltonian mechanics.
SIAM Journal on Numerical Analysis, 28(5):1463–1516, 1991.

Gene H. Golub and Charles F. van Loan.
Matrix computations.
Johns Hopkins Press, Baltimore, 3rd edition, 1996.

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

Zhongxiao Jia.
The regularization theory of the Krylov iterative solvers LSQR and CGLS for linear discrete ill-posed problems, part I: the simple singular value case.
https://arxiv.org/abs/1701.05708, January 2017.

Rasmus Larsen.
Lanczos bidiagonalization with partial reorthogonalization.
DAIMI Report Series, 27(537), Dec. 1998.

B. J. Last and K. Kubik.
Compact gravity inversion.
Geophysics, 48(6):713–721, 1983.

Christopher C. Paige and Michael A. Saunders.
LSQR: an algorithm for sparse linear equations and sparse least squares.
ACM Trans. Math. Software, 8(1):43–71, 1982.

Y. Saad.

Numerical Methods for Large Eigenvalue Problems.
Society for Industrial and Applied Mathematics, 2011.

A. Saibaba.
Randomized subspace iteration: Analysis of canonical angles and unitarily invariant norms.
SIAM Journal on Matrix Analysis and Applications, 40(1):23–48, 2019.

Saeed Vatankhah, Rosemary A. Renaut, and Vahid E. Ardestani.
A fast algorithm for regularized focused 3-D inversion of gravity data using the randomized SVD.
Geophysics, 2018.

Brendt Wohlberg and Paul Rodríguez.
An iteratively reweighted norm algorithm for minimization of total variation functionals.
Signal Processing Letters, IEEE, 14(12):948–951, 2007.