Examining the Feasibility of Computational Methods for Inverting Gravity and Magnetic Data: Randomized or Krylov Factorizations?

Rosemary Renaut\textsuperscript{1}, Jarom Hogue\textsuperscript{1} and Saeed Vatankhah\textsuperscript{2}

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

\textsuperscript{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

June 24 2021
Outline

Motivation and Background
  Inversion Under-sampled Magnetic / Gravity Data
  Basic technique: the singular value decomposition (SVD)

Approximating the SVD
  Krylov: Golub Kahan Bidiagonalization - LSQR
  Randomized Singular Value Decomposition

Motivating Example on RSVD and GKB
  Computational Costs to Convergence

Numerical Experiments

Conclusions and Practical Data
The Fredholm Integral of the First Kind

\[ d(a, b, c) = \int \int \int h(a, b, c, x, y, z) \zeta(x, y, z) dx \, dy \, dz \]

Linear Model

- Integral equation describes a convolution with
  \[ h(a, b, c, x, y, z) = h(x - a, y - b, z - c) \] spatially invariant.
- \[ \zeta(x, y, z) \] gives the unknown parameter on the volume domain, discretized as vector \( \mathbf{m} \)
- \( d \) response at the surface, \( c = 0 \), discretized as \( \mathbf{d} \).
- Matrix \( G \) discretizes \( h \) integral over volume cells.
- \( G \): block Toeplitz with Toeplitz blocks (BTTB) by depth.
- Fast evaluation \( G \mathbf{m} \) and \( G^T \mathbf{d} \) by Fast Fourier Transform.
Two Geophysics Examples: Solve $W_{d_i} G_i(W_{\text{depth}})_i m_i \approx W_{d_i} d_i$

<table>
<thead>
<tr>
<th>Gravity Problem</th>
<th>Magnetic Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_1 \cong G_1 m_1$</td>
<td>$d_2 \cong G_2 m_2$</td>
</tr>
<tr>
<td>$G_1$ is symmetric ($m \times n$)</td>
<td>$G_2$ is asymmetric ($m \times n$)</td>
</tr>
<tr>
<td>$d_1$ gravity anomaly ($m$)</td>
<td>$d_2$ magnetic anomaly ($m$)</td>
</tr>
<tr>
<td>$m_1$ density ($n$)</td>
<td>$m_2$ susceptibility ($n$)</td>
</tr>
</tbody>
</table>

**Common Assumptions:** (also to include a priori geophysical information: Cassiani Session I)

- $n \gg m$, under sampled and $G_i$ ill-conditioned
- Depth weighting on $m_i$: diagonal ($W_{\text{depth}})_i$ counteracts natural decay of kernels.
- Data are weighted by inverse covariance of measurement noise, independent. $W_{d_i}$
- Stabilized Inversion to find $m_i$ given $d_i$ is required. (de Alba Session I)
- Often regularization parameters are fixed.
Standard Form Regularization

\[ x(\alpha) = \arg\min_x \{ \| b - Ax \|^2 + \alpha^2 \| x \|^2_2 \} \]

Singular value decomposition (SVD) Expansion Filters \( \gamma_i \).

\[
x = \sum_{i=1}^{r} \gamma_i(\alpha) \begin{bmatrix} s_i \\ \sigma_i \end{bmatrix} v_i, \quad \gamma_i(\alpha) = \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2}, \ i = 1 \ldots r,
\]

Truncated SVD \( A_\ell = U_\ell \Sigma_\ell V_\ell^\top \)

Truncated solution \( \alpha \) found by standard method

\[
x = V_\ell \Gamma(\alpha, \Sigma_\ell) U_\ell^\top b \quad \gamma_i(\alpha) = \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2}, \ i = 1 \ldots \ell,
\]

Solves Standard Form

\[ x_\ell(\alpha) \approx \arg\min_x \{ \| b - A_\ell x \|^2 + \alpha_\ell^2 \| x \|^2_2 \} \]
Generalized Tikhonov - with prior information $x_{\text{prior}}$ and operator $L$

$$x(\alpha) = \arg\min_x \{ \| b - Ax \|^2 + \alpha^2 \| L(x - x_{\text{prior}}) \|^2 \}$$

$L$ is invertible  Rewrite using $y = L(x - x_{\text{prior}})$

$$d = b - Ax_{\text{prior}}$$

$$y(\alpha) = \arg\min_y \{ \| d - AL^{-1}y \|^2 + \alpha^2 \| y \|^2 \}$$

Truncate  the SVD for $(AL^{-1})_\ell$

$$x_\ell(\alpha) \approx x_{\text{prior}} + L^{-1}(\arg\min_y \{ \| d - (AL^{-1})_\ell y \|^2 + \alpha_\ell^2 \| y \|^2 \})$$

Advantage of SVD  Find optimal $\alpha_\ell$ by standard techniques.

Can these techniques be employed for large scale inversion?
Applying Generalized Regularization for Focusing Inversion

Replace $L$ by $W_L^{(k)}$ iteration $k$: Iteratively reweighted LS (IRLS)

$$(W_L^{(k)})_{ii} = ((x_i^{(k-1)} - x_i^{(k-2)})^2 + \epsilon)^{(p-2)/4} \quad \epsilon > 0$$

$p = 1 : \|W_L^{(k)}x\|_2^2 \simeq \|x\|_1$. ($p = 2$ yields standard $\|x\|_2$.)

Invertibility $\epsilon > 0$ [WR07].

Right preconditioner : $(W^{(k)})^{-1}$ is a right preconditioner for $A$

Apply stabilized inversion for approximate truncated SVD With updating right hand side $b = b^{(k)}$, update the system matrix

$$A^{(k)}_\ell \simeq W_dGW^{-1}_{\text{depth}}(W_L^{(k)})^{-1}$$

Regularization Parameter Estimate $\alpha^{(k)}_\ell$ each iteration.

How to obtain $A^{(k)}_\ell$ for the large scale?
Obtaining an Approximate SVD and Regularizing

- Factorization by GKB: $AG_\ell = H_{\ell+1}B_\ell$
- Find the SVD of $B_\ell$: $B_\ell = \tilde{U}_\ell \tilde{\Sigma}_\ell \tilde{V}_\ell^\top$
- Approximate SVD for $A$: $A_\ell \cong (H_{\ell+1}\tilde{U}_\ell)\tilde{\Sigma}_\ell (\tilde{V}_\ell^\top G_\ell^\top)$
- Projected GKB Problem for Focusing Inversion

$$w^{(k)}_\ell (\zeta^{(k)}_\ell) = \arg\min_{w \in \mathcal{R}_\ell}\{ \| B^{(k)}_\ell w - b^{(k)} \|_2 e_1^{(\ell+1)} \frac{2}{2} + (\zeta^{(k)}_\ell)^2 \| w \|_2^2 \}.$$

- Given $\tilde{U}_\ell \tilde{\Sigma}_\ell \tilde{V}_\ell^\top$ use standard technique to find $\zeta^{(k)}_\ell$
- $y^{(k)} = G_{\ell p} w^{(k)}_\ell (\zeta^{(k)}_\ell)$
Randomized Singular Value Decomposition [HMT11, VRA18]

Obtaining an Approximate SVD and Regularizing

- Generate a random matrix \( \Omega \in \mathbb{R}^{\ell_p \times m} \).
- Let \( Y = \Omega A \in \mathbb{R}^{\ell_p \times n} \).
- \( \text{Range}(Y^T) \) approximates dominant range(\( G^T \)).
- \( \text{Orth}(Y^T) = Q \) approximates right singular space of \( G \).
- \( B = AQ \in \mathbb{R}^{m \times \ell_p} \) projects left singular space to \( B \).
- For small \( \ell \) \( B \) is well conditioned.
- Eigen decomposition \( B^T B = [\bar{V}, D] \) calculated cheaply.
- Obtain \( A \approx U_{\ell} \bar{\Sigma}_\ell \bar{V}_{\ell}^T \).
- Projected RSVD Problem for Focusing Inversion

\[
x^{(k)}_{\ell} (\mu^{(k)}_{\ell}) = \arg\min_{x \in \mathbb{R}^\ell} \left\{ \| A^{(k)}_{\ell} x - b^{(k)} \|_2^2 + (\mu^{(k)}_{\ell})^2 \| x \|_2^2 \right\}.
\]
## Summary Comparisons: \( \ell \) approximation of \( A \)

<table>
<thead>
<tr>
<th>Model</th>
<th>TSVD</th>
<th>LSQR</th>
<th>RSVD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVD</td>
<td>( A_\ell ) ( U_\ell \Sigma_\ell V_\ell^\top )</td>
<td>( \tilde{A}<em>\ell ) ( (H</em>{\ell+1} \tilde{U}<em>\ell) \tilde{\Sigma}</em>\ell (G_\ell \tilde{V}_\ell)^\top )</td>
<td>( \overline{A}<em>\ell ) ( U</em>\ell \tilde{\Sigma}<em>\ell \tilde{V}</em>\ell^\top )</td>
</tr>
<tr>
<td>Terms</td>
<td>( u_i^\top b )</td>
<td>( (H_{\ell+1} \tilde{U}_\ell)_i^\top b )</td>
<td>( \tilde{u}_i^\top b )</td>
</tr>
<tr>
<td>Basis</td>
<td>( v_i )</td>
<td>( (G_\ell \tilde{V}_\ell)_i )</td>
<td>( \tilde{v}_i )</td>
</tr>
<tr>
<td>Coeff</td>
<td>( \gamma_i(\alpha_\ell \frac{s_i}{\sigma_i}) v_i )</td>
<td>( \gamma_i(\zeta_\ell \frac{s_i}{\sigma_i})(G_\ell \tilde{V}_\ell)_i )</td>
<td>( \gamma_i(\mu_\ell \frac{s_i}{\sigma_i}) \tilde{v}_i )</td>
</tr>
<tr>
<td>Condition</td>
<td>Depends on ( \sigma_{\ell+1} )</td>
<td>Depends on ( A )</td>
<td>Depends on ( A_\ell )</td>
</tr>
</tbody>
</table>

**Properties depend on space \( \ell \) and oversampling \( p \)**

**Given a TSVD we can find** \( \alpha_\ell, \mu_\ell, \) or \( \zeta_\ell \)**
Example Magnetic for subspace size $\ell m = 5000, n = 75000$ SNR: 19

Figure: Model

Figure: True Data

Figure: Noisy Data

Figure: Krylov $k = 15, 1s$

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

(For Gravity RSVD is faster, [VRA17], here GKB is faster)
Computational Cost Estimates per iteration

**Comparison in terms of flops**

*2DFFT* for matrix operations: Approximate rank $\ell$ SVD

\[
\text{Cost}_{\text{GKB}} = 8n\ell \log_2(4m) + 2\ell^2(n + m) + \text{LOT} \\
\text{Cost}_{\text{RSVD}} = 16n\ell \log_2(4m) + 4\ell^2(2n + 3/2m) + 5\ell^3 + \text{LOT}.
\]

For same size $\ell$:

\[
\text{Cost}_{\text{RSVD}} \approx 2\text{Cost}_{\text{GKB}}
\]

Equivalently, require

- dominant space $\ell$ for RSVD to be one half size for GKB with same number of iterations
- or convergence of IRLS in one half number of operations

Suggests different properties of magnetic and gravity if performance differences
Relative Costs GKB:RSVD for Gravity and Magnetic: Increasing $m$ to convergence and increasing $\ell$

**Gravity**

$\frac{\text{Cost2DFFT}_{\text{GKB}}}{\text{Cost2DFFT}_{\text{RSVD}}}$

**Magnetic**

$\frac{\text{Cost2DFFT}_{\text{GKB}}}{\text{Cost2DFFT}_{\text{RSVD}}}$

The green horizontal line is at $y = 1$.

**Comparison: approximation to $A_\ell$ to convergence of IRLS**

- GKB is more expensive for gravity problem
- RSVD is more expensive for magnetic problem unless $\ell$ close to $m$
Inversion of Gravity Data for increasing $m$: $\ell = \text{floor}(m/8)$

**Figure:** GKB: $m = 6000$, $\ell = 750$, $(K, Time, RE) = (9, 248s, .76)$

![Contour map for GKB with m = 6000, l = 750](image1)

**Figure:** GKB: $m = 18375$, $\ell = 2296$, $(K, Time, RE) = (11, 5732s, .75)$

![Contour map for GKB with m = 18375, l = 2296](image2)

**Figure:** RSVD: $m = 6000$, $\ell = 750$, $(K, Time, RE) = (6, 216s, .57)$

![Contour map for RSVD with m = 6000, l = 750](image3)

**Figure:** RSVD: $m = 18375$, $\ell = 2296$, $(K, Time, RE) = (7, 3311s, .60)$

![Contour map for RSVD with m = 18375, l = 2296](image4)
Inversion of Magnetic Data for increasing \( m \): \( \ell = \text{floor}(m/8) \)

**Figure:** GKB: \( m = 6000, \ell = 750 \), 
\( (K, Time, RE) = (5, 136s, .63) \)

**Figure:** GKB: \( m = 18375, \ell = 2296 \), 
\( (K, Time, RE) = (7, 3809s, .67) \)

**Figure:** RSVD: \( m = 6000, \ell = 750 \), 
\( (K, Time, RE) = (25, 883s, .64) \)

**Figure:** RSVD: \( m = 18375, \ell = 2296 \), 
\( (K, Time, RE) = (14, 6618s, .70) \)
Conclusion: It is feasible to do automated focusing inversion

1. The use of the BTTB structure enables detailed analysis for large scale resolutions
2. RSVD and GKB can be used in the context of focusing inversion of gravity and magnetic potential field data
3. It is sufficient to find low rank SVD approximations for the weighted sensitivity matrices with $\ell \approx \frac{m}{8}$
4. Use of the low rank SVD permits use of automated estimation of regularization parameters
5. GKB is preferred to invert magnetic data sets. (Decay rate of $\sigma_i$ is mild)
6. RSVD is preferred to invert gravity data sets. (Decay rate of $\sigma_i$ is moderate)
THANK YOU
QUESTIONS
Application for Real Data, $m = 3184$ measurements, padded to 5184 using $\ell = 480 \approx m/8$, oversampled $\ell_p = 504$. [Pil09, VLR$^+$20, RHVL20]
Results Isosurface

Figure: The isosurface

- 200 depth levels
- Convergence in 18 iterations
- 3690s on MacBook Pro Laptop with 2.5GHz Dual Core and 16GB memory.
Cross Sections

**Figure: 300m**

**Figure: 600m**

**Figure: 900m**
Some key references

N. Halko, P. G. Martinsson, and J. A. Tropp.
Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. 

Mark Pilkington.
3D magnetic data-space inversion with sparseness constraints. 

Christopher C. Paige and Michael A. Saunders.
LSQR: an algorithm for sparse linear equations and sparse least squares. 

Rosemary A. Renaut, Jarom D Hogue, Saeed Vatankhah, and Shuang Liu.
A fast methodology for large-scale focusing inversion of gravity and magnetic data using the structured model matrix and the 2D fast Fourier transform. 

Saeed Vatankhah, Shuang Liu, Rosemary Anne Renaut, Xiangyun Hu, and Jamaledin Baniamerian.
Improving the use of the randomized singular value decomposition for the inversion of gravity and magnetic data. 

Saeed Vatankhah, Rosemary A. Renaut, and Vahid E. Ardestani.
3-D projected $\ell_1$ inversion of gravity data using truncated unbiased predictive risk estimator for regularization parameter estimation. 

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. 

Brendt Wohlberg and Paul Rodríguez.
An iteratively reweighted norm algorithm for minimization of total variation functionals. 
GKB: The Golub Kahan Bidiagonalization algorithm to obtain solution $y$

Input $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$, a target rank $\ell$ and size of oversampled projected problem $\ell_p$, $\ell < \ell_p \ll m$.

1. Set $a = \text{zeros}(n, 1)$, $B = \text{sparse}(\text{zeros}(\ell_p + 1, \ell_p))$, $H = \text{zeros}(m, \ell_p + 1)$, $G = \text{zeros}(n, \ell_p)$

2. Set $\beta = \|b\|_2$, $h = b/\beta$, $H(:, 1) = h$

3. For $i = 1 : \ell_p$
   3.1 $b = A^\top h - \beta a$
   3.2 For $j = 1 : i - 1$ \{ $b = b - (G(:, j)^\top b)G(:, j)$ (modified Gram-Schmidt (MGS)) \}
   3.3 $\gamma = \|b\|_2$, $a = b/\gamma$, $B(i, i) = \gamma$, $G(:, i) = a$
   3.4 $c = Aa - \gamma h$
   3.5 For $j = 1 : i$ \{ $c = c - (H(:, j)^\top c)H(:, j)$ (MGS) \}
   3.6 $\beta = \|c\|_2$, $h = c/\beta$, $B(i + 1, i) = \beta$, $H(:, i + 1) = h$

4. SVD for sparse matrix: $\tilde{U}_{\ell_p} \tilde{\Sigma}_{\ell_p} \tilde{V}_{\ell_p}^\top = \text{svds}(B, \ell_p)$

5. Find regularization $\zeta_{\ell}$ using $\tilde{U}_{\ell_p}(:, 1 : \ell)$ and $\tilde{\Sigma}_{\ell_p}(1 : \ell, 1 : \ell)$

6. Solution $y = \|b^{(k)}\|_2 (G_{\ell_p} \tilde{V}_{\ell_p}) \Gamma(\zeta_{\ell}, \tilde{\Sigma}_{\ell_p}) \tilde{U}_{\ell_p} (1, :)^\top$
Finding the $y$ using Randomization for the SVD [HMT11]

Input $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$, target rank $\ell$, power iteration $q$ and oversampling $\ell_p$, $\ell < \ell_p \ll m$.

1. Find Gaussian random $\Omega \in \mathbb{R}^{\ell_p \times m}$, $Y = \Omega A \in \mathbb{R}^{\ell_p \times n}$

2. Orthogonalize $[Q, \sim] = qr(Y^T, 0)$, $Q \in \mathbb{R}^{n \times \ell_p}$.

3. If $q > 0$ Repeat $q$ times:
   3.1 $Y = AQ \in \mathbb{R}^{m \times \ell_p}$. $[Q, \sim] = qr(Y, 0)$, $Q \in \mathbb{R}^{m \times \ell_p}$.
      $Y = Q^\top A$, $Y \in \mathbb{R}^{\ell_p \times n}$, $[Q, \sim] = qr(Y^\top, 0)$, $Q \in \mathbb{R}^{n \times \ell_p}$

4. $B = AQ \in \mathbb{R}^{m \times \ell_p}$ and $Y = B^\top B \in \mathbb{R}^{\ell_p \times \ell_p}$

5. Eigen-decomposition of $B^\top B$: $[\overline{V}, D] = \text{eig}((Y + Y^\top)/2)$

6. Form the SVD:
   6.1 $[S, is] = \text{diag}(\sqrt{\text{sort}(|\text{real}(D)|,'\text{descend}')})$.
   6.2 $\overline{\Sigma}_\ell = \text{diag}(S(1 : \ell)) \overline{V} = \overline{V}(::, is(1 : \ell)) \overline{U} = \overline{V}./ (S(1 : \ell)^\top)$

7. Find regularization $\mu_\ell$ using $\overline{U}$, $\overline{\Sigma}_\ell$, and $B^\top b$

8. Solution $y = (Q \overline{V}) \Gamma(\mu_\ell, \overline{\Sigma}_\ell) \overline{U}^\top (B^\top b)$