Off-the-Grid Low-Rank Matrix Recovery: Seismic Data Reconstruction

Oscar López

2016 Canadian Mathematical Society Summer Meeting

June 30, 2016

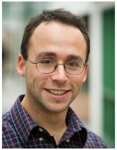
Contributors

University of Brisith Columbia
Seismic Laboratory for Imaging and Modeling (and friends)


R. Kumar


O. Yilmaz


F.J. Herrmann


A. Arakvin


H. Mansour


B. Recht

Quick Summary

- **Problem**: large-scale seismic data interpolation.

Quick Summary

- **Problem**: large-scale seismic data interpolation.
- **Approach**: organize seismic data in low-rank matrix form and exploit this low-dimensional structure.

Quick Summary

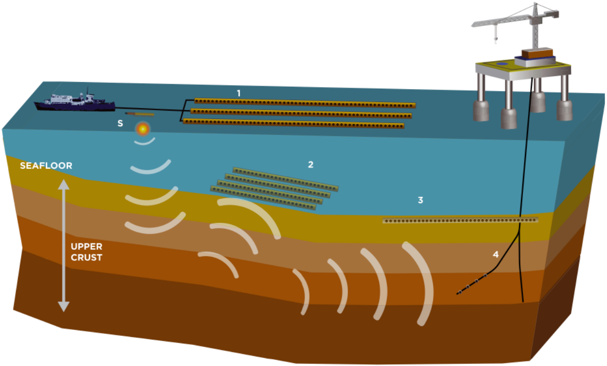
- **Problem**: large-scale seismic data interpolation.
- **Approach**: organize seismic data in low-rank matrix form and exploit this low-dimensional structure.
- **Outcome**: frugal interpolation procedure taylored for large data sets.

Table of Contents

Seismic Data Interpolation

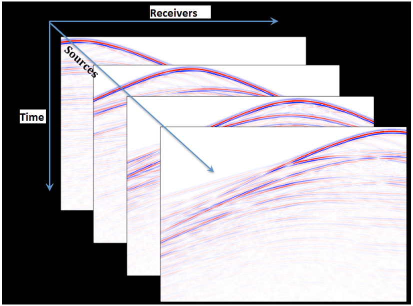
Seismic Data Acquisition

Sound waves are sent into the earth.
Sensors record the strength and travel time of these waves through the layers in the earth's crust.



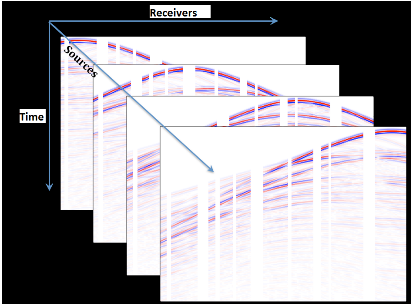
2D Seismic Data

For a 2D image of the subsurface we obtain data cube $\mathcal{D} \in \mathbb{R}^{n_t \times n_r \times n_s}$



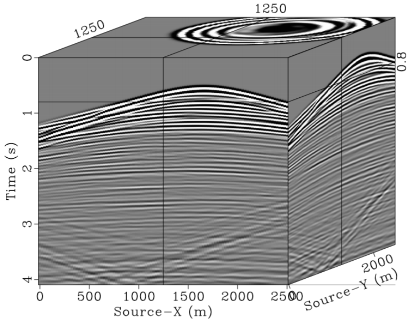
2D Seismic Data

Since dense acquisition is expensive, several sources and/or receivers are removed.



3D Seismic Data

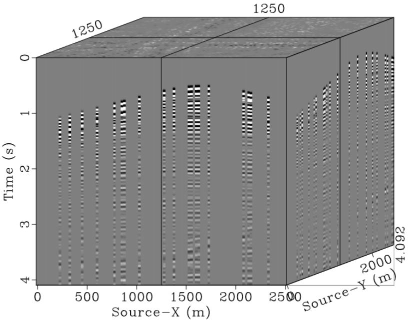
For a 3D image of the subsurface we obtain a 5D data set $\mathcal{D} \in \mathbb{R}^{n_t \times n_{rx} \times n_{ry} \times n_{sx} \times n_{sy}}$



Common Receiver Gather, fix (receiver-x, receiver-y)

3D Seismic Data

Since dense acquisition is expensive, several sources and/or receivers are removed.



Subsampled Common Receiver Gather

Seismic Data Interpolation

Infer 3D velocity model from multi-experiment data:

- $\mathcal{O}(10^{15})$ datapoints
- Subsampled data needs to be interpolated for accurate results.
- Typical acquisition/sampling schemes reach a bottleneck in subsampling ratio.
  Need new methodology to lessen data deluge.

Low-Rank Matrix Recovery

Low-Rank Matrix Recovery

Goal: Recover matrix $M \in \mathbb{C}^{n \times m}$ from $k \ll nm$ noisy linear measurements

$$y_i = \langle M, \mathcal{A}_i \rangle + d_i, \quad 1 \leq i \leq k.$$

measurement ensemble: $\mathcal{A} : \mathbb{C}^{n \times m} \mapsto \mathbb{C}^k$
noise model: $d \in \mathbb{C}^k$, $\|d\|_2 \leq \eta$

Low-Rank Matrix Recovery

Goal: Recover matrix $M \in \mathbb{C}^{n \times m}$ from $k \ll nm$ noisy linear measurements

$$y_i = \langle M, \mathcal{A}_i \rangle + d_i, \quad 1 \leq i \leq k.$$

Key Assumption: M has rank $r \ll \min(n, m)$, or can be well approximated by such a "low-rank" matrix.

Low-Rank Matrix Recovery

Approximate matrix as $M^\sharp$, where

$$M^\sharp := \operatorname*{arg\,min}_{X \in \mathbb{C}^{n\times m}} \ \operatorname{rank}(X) \ \text{ s.t. } \ \|\mathcal{A}(X) - y\|_2 \leq \eta.$$

Low-Rank Matrix Recovery

Approximate matrix as $M^\sharp$, where

$$M^\sharp := \operatorname*{arg\,min}_{X \in \mathbb{C}^{n \times m}} \ \text{rank}(X) \ \text{ s.t. } \ \|\mathcal{A}(X) - y\|_2 \leq \eta.$$

This problem is NP-hard in general.

Nuclear Norm Minimization

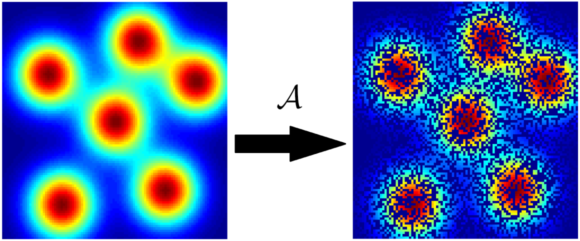
Instead, approximate rank by convex surrogate

$$M^\sharp := \operatorname*{arg\,min}_{X \in \mathbb{C}^{n \times m}} \|X\|_* \;\; \text{s.t.} \;\; \|\mathcal{A}(X) - y\|_2 \leq \eta,$$

where $\|X\|_* = \sum_{\ell=1}^{\min(n,m)} \sigma_\ell(X)$ denotes the nuclear norm of X.

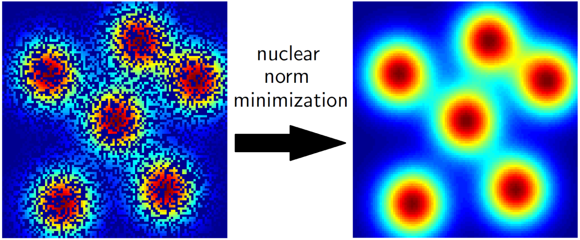
Example: Matrix Completion

Observations: Subset of matrix entries



Example: Matrix Completion

Typical Results: Can recover "incoherent" matrix from $k \sim \mathcal{O}(nr\log^2(n))$ entries chosen randomly.



B. Recht [2011]. A Simpler Approach to Matrix Completion. *The Journal of Machine Learning Research*.

Low-Rank Matrix Recovery

Can consider the seismic data interpolation problem as an instance of matrix completion.

Potential reduction in the number of measurements required in acquisition (proportional to rank $r \ll \min(n_s, n_r)$).

Challenges

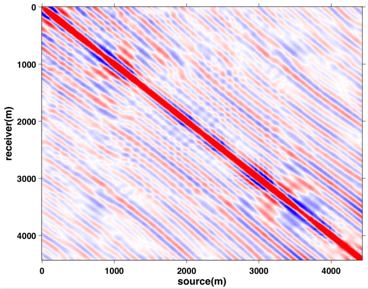
Challenge 1 - Low-Rank Structure of Seismic Data

Simply considering standard seismic data volumes does not immediately yield low-rank structure.

Need to find transform domain where seismic data matrix exhibits quickly decaying singular values.

Acquisition Domain

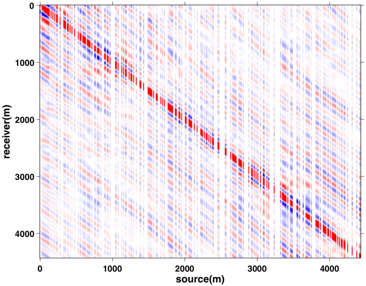
Apply Fourier transform along the time axis, consider a monochromatic frequency slice: $M \in \mathbb{C}^{n_s \times n_r}$



Not "low-rank" due to strong diagonal energy.

Acquisition Domain

Subsampled data:
$\mathcal{A}(M) \in \mathbb{C}^{n_s \times n_r}$



Missing sources (and/or receivers) can only **decrease rank**.

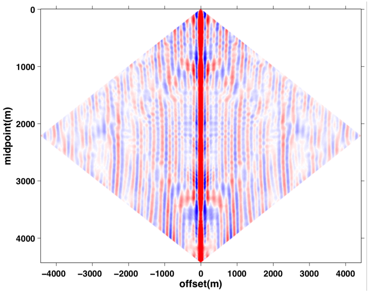
Midpoint-Offset Domain

Midpoint-offset coordinates:
midpt$= \frac{s+r}{2}$, offset$= \frac{s-r}{2}$

Midpoint-offset domain transform operator:
$\mathcal{M} : \mathbb{C}^{n\times m} \mapsto \mathbb{C}^{n_{mid}\times m_{off}}$

Midpoint-Offset Domain

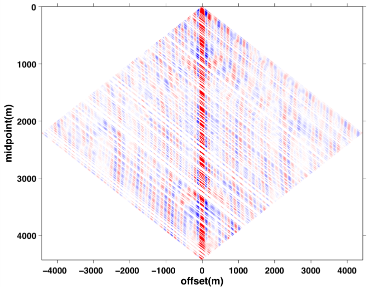
Matrix in transform domain: $\mathcal{M}(M) \in \mathbb{C}^{n_{mid} \times m_{off}}$



Exhibits **“low-rank” structure**, due to source-receiver reciprocity and vertical alignment of wavefronts.

Acquisition Domain

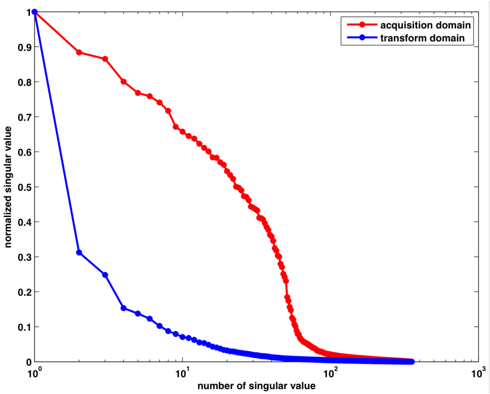
Subsampled matrix in transform domain: $\mathcal{M}\mathcal{A}(M) \in \mathbb{C}^{n_s \times n_r}$



Missing sources (and/or receivers) correspond to missing diagonals, which **do** affect the low-rank structure.

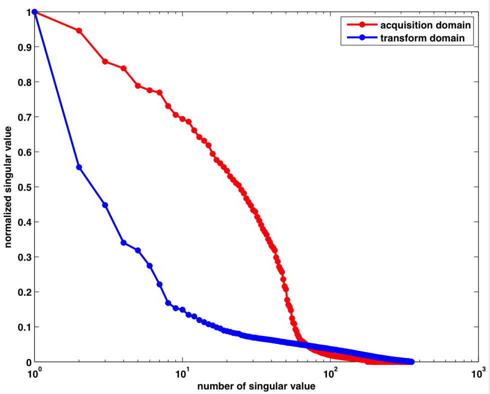
Low-Rank Structure Of Seismic Data

Decay of singular values of **fully sampled** matrices



Low-Rank Structure Of Seismic Data

Decay of singular values of **subsampled** matrices



Low-Rank Structure of Seismic data

Quickly decaying singular values in transform domain (i.e., "low-rank structure"). Missing entries affect the low-rank structure.

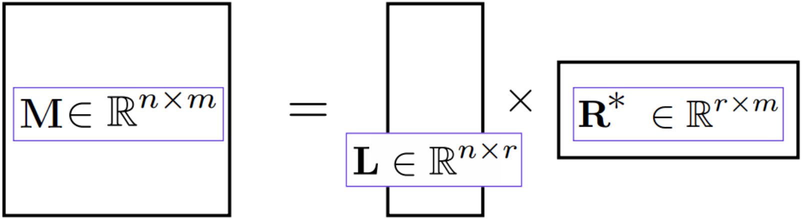
Interpolate sequentially along the frequency axis in transform domain.

Challenge 2 - Big Data

Methods for nuclear norm minimization require SVD computations at each iteration (e.g., singular value projection and singular value thresholding).

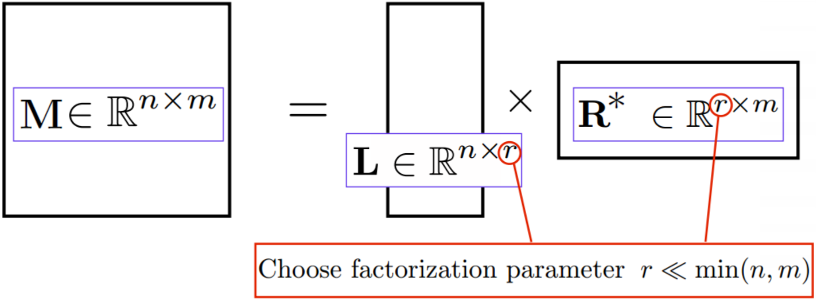
Such computations are prohibitely expensive for large matrices: $M \in \mathbb{C}^{4000\times 4000}$.

Factorization Approach



$$M = LR^*$$

Factorization Approach



$$M = LR^*$$

LR-BPDN

Since

$$\|M\|_* = \min_{M=LR^*} \quad \frac{1}{2}\left(\|L\|_F^2 + \|R\|_F^2\right)$$

N. Srebro, et al. [2005]. Maximum-Margin Matrix Factorization. *Advances in Neural Information Processing Systems*.

we can instead solve

$$\min_{L\in\mathbb{C}^{n\times r}, R\in\mathbb{C}^{m\times r}} \quad \frac{1}{2}\left(\|L\|_F^2 + \|R\|_F^2\right) \quad \text{s.t.} \quad \|\mathcal{A}(LR^*) - y\|_2 \leq \eta,$$

LR-BPDN

- Reduced memory requirements from nm to $nr + mr$.
- No SVD computations.
- Many existing methods to handle bi-convex optimization.

Challenge 3 - Sampling Off-the-Grid

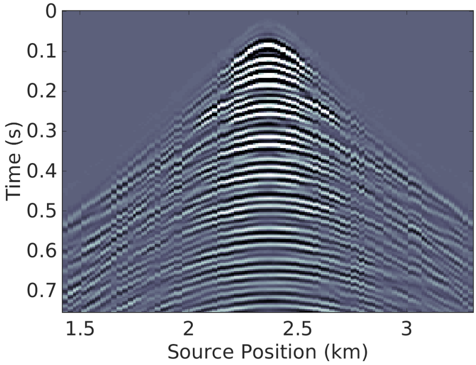
Due to fallible physical conditions in seismic surveys, sources and receivers are often deviated from a regularly spaced sampling grid.

Such irregular data matrices are less fit for the low-rank model, and therefore subject to degraded reconstruction.

Challenge 3 - Sampling Off-the-Grid

.

Irregular common receiver gather: deviated source positions



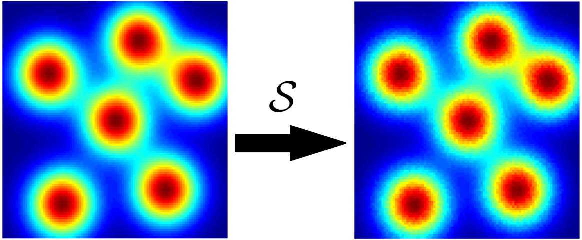
Regularization Operator

Introduce regularization operator: simulates effect of unstructed grid
$\mathcal{S} = \mathcal{N}\mathcal{F}^* : \mathbb{C}^{n\times m} \mapsto \mathbb{C}^{n\times m}$

$\mathcal{F}$ is a 2D fast Fourier transform
$\mathcal{N}$ is 2D non-uniform fast Fourier transform, whose frequencies are chosen according to known sampling locations

Regularization Operator

Regularization operator maps regular data to irregular data



Improved measurement model, $y = \mathcal{A}\mathcal{S}(M)$

Regularization Operator

- Accurate for Lipschitz continuous signals. Error bound, ϵ, depends on Lipschitz constant, n, m and extent of grid deviations.
- $\mathcal{AS}$ preserves restricted isometry constant of subgaussian measurement operators, $\mathcal{A}$.

Restricted Isometry Constant

Definition

For a linear map $\mathcal{A} : \mathbb{C}^{n\times m} \mapsto \mathbb{C}^k$ and $r \leq \min\{n, m\}$, the rank-r restricted isometry constant $\delta_r(\mathcal{A})$ is defined to be the smallest $0 < \delta < 1$ such that

$$(1-\delta)\|X\|_F^2 \leq \|\mathcal{A}(X)\|_2^2 \leq (1+\delta)\|X\|_F^2$$

for all $X \in \mathbb{C}^{n\times m}$ of rank at most r.

Restricted Isometry Constant: $\mathcal{AS}$

Theorem

Let $\mathcal{A}$ be a normalized subgaussian measurement map with $\max_{p,k,\ell} \|\mathcal{A}(p,k,\ell)\|_{\psi_2} = K$. Define $\mathcal{S}$ according to randomly perturbed sampling locations, i.i.d. with any distribution. Then for any $0 < \delta \leq 1$, $0 < \gamma < \frac{1}{2}$ and rank r matrix, $X \in \mathbb{C}^{n \times m}$,

$$\left(1 - \frac{\delta}{1-2\gamma}\right)\|X\|_F^2 \leq \|\mathcal{AS}(X)\|_2^2 \leq \left(1 + \frac{\delta}{1-2\gamma}\right)\|X\|_F^2$$

with probability at least $1-\tau$ provided that

$$N \geq \frac{4K^4}{\bar{c}\delta^2}\left(r(m+n+1)\log\left(1+\frac{6}{\gamma}\right) + \log(2\tau^{-1})\right),$$

where $\bar{c} > 0$ is an absolute constant.

Final Product

$$(L^\sharp, R^\sharp) := \underset{L\in\mathbb{C}^{n\times r}, R\in\mathbb{C}^{m\times r}}{\arg\min} \quad \frac{1}{2}\left(\|L\|_F^2 + \|R\|_F^2\right)$$
$$\text{s.t.} \quad \|\mathcal{A}\mathcal{S}\mathcal{M}^*(LR^*) - y\|_2 \leq \eta + \epsilon,$$

$$M^\sharp = \mathcal{M}^*(L^\sharp R^{\sharp *})$$

Nuclear Norm Minimization

Theorem

Let $M^\sharp$ be the solution of the nuclear norm minimization problem. If

$$\delta_{2r}(\mathcal{A}\mathcal{S}) < \frac{4}{\sqrt{41}}$$

then we have

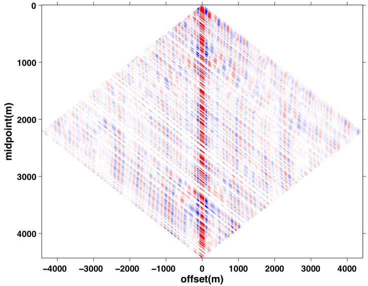
$$\|M - M^\sharp\|_F \leq \frac{c_1}{\sqrt{r}} \sum_{\ell=r+1}^{\min\{n,m\}} \sigma_\ell(\mathcal{M}(M)) + c_2\eta + c_3\epsilon,$$

where $c_1, c_2, c_3 > 0$ are constants depending only on $\delta_{2r}(\mathcal{A})$.

Experiments

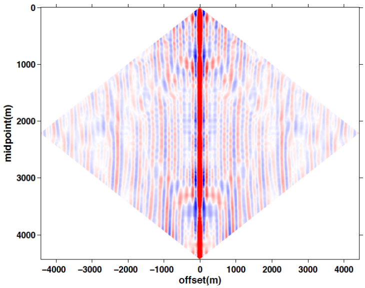
10 Hz frequency slice - 354 × 354 matrix

75% randomly missing sources
subsampled data in transform domain



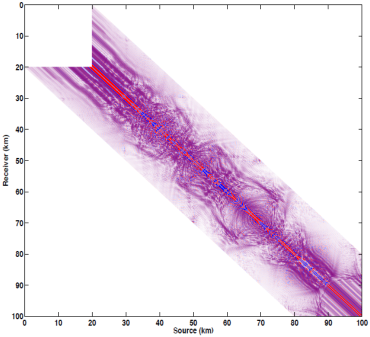
10 Hz frequency slice - 354 × 354 matrix

transform domain recovery
SNR = 18.5 dB



Synthetic, Gulf of Mexico

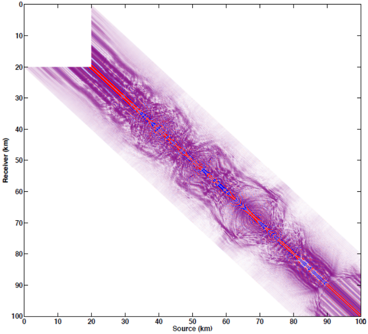
4001×4001 matrix, at 7 Hz



Recover from 80% randomly missing sources

Synthetic, Gulf of Mexico

Recovery, SNR = 14.2 dB

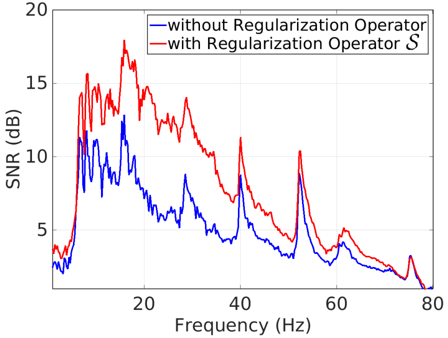


time = 20 minutes

Gulf of Suez Irregular Data Cube

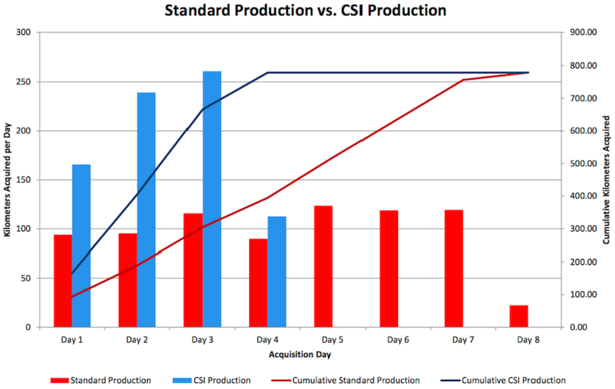
$\mathcal{D} \in \mathbb{R}^{1024 \times 354 \times 354}$

Compare recovery with vs without regularization operator, $\mathcal{S}$.



Compressive Sensing in the field

ConocoPhilips successfully applied CS in land and marine settings



Thanks to Chuck Mosher.

Conclusions and On-Going work

Seismic data exhibits low-rank structure that can be exploited.

Bi-convex optimization seems to get the job done.

On-Going Work:

- How do we choose factorization parameter, r?
- How should we subsample in a practical maner?
- 3D data. Matricize or deal with tensor?
- Alternate optimization over matrix factors.
- What if deviated source/receiver positions are not known?

Acknowledgements

Thank you for your attention!

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BGP, CGG, Chevron, ConocoPhillips, DownUnder GeoSolutions, Hess Corporation, Petrobras, PGS, Sub Salt Solutions, WesternGeco, and Woodside. O. Ylmaz also acknowledges an NSERC Discovery Grant (22R82411) and an NSERC Accelerator Award (22R68054).