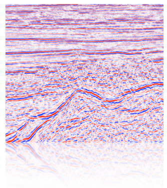
A primer on stable signal recovery

Gilles Hennenfent
PhD student

Felix Herrmann
Assistant professor

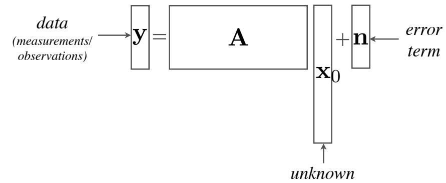
Seismic Laboratory for Imaging & Modeling
Department of Earth & Ocean Sciences
The University of British Columbia



SINBAD meeting: General introduction session
August 28, 2006

Problem statement

consider the following (severely) underdetermined system of linear equations



is it possible to recover $\mathbf{x}_0$ accurately from $\mathbf{y}$?

- general answer is NO...
  - infinitely many solutions

$$\begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \\ 0 \\ 0 \end{bmatrix} \quad \text{but} \quad \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 50 \\ 47 \\ 0 \end{bmatrix}$$

desired solution — other possible solution

  - least squares solution

$$\mathbf{x}_{LS} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y} \qquad \mathbf{x}_{LS} = \begin{bmatrix} 0 \\ 1.5 \\ -1.5 \\ 0 \end{bmatrix}$$

- BUT, under some specific conditions on the matrix **A** and the solution $\mathbf{x}_0$, the answer is YES!

References

- E. Candès, J. Romberg, "*Practical signal recovery from random projections*", IEEE Trans. Signal Processing, 2005.
- E. Candès, J. Romberg, T. Tao, "*Stable signal recovery from incomplete and inaccurate measurements*", Comm. on Pure and Applied Math, 2005.
- J. Romberg, "*Sparse signal recovery via l_1 minimization* ", 2006.
- E. Candès, T. Tao, "*Decoding by linear programming*", IEEE Transactions on information theory, vol. 51, no. 12, December 2005, pp. 4203- 4215.
- D. L. Donoho, "*Compressed sensing*", IEEE transactions on information theory, vol. 52, no. 4, April 2006, pp. 1289-1306.
- Y. Tsaig, D. L. Donoho, "*Extensions of compressed sensing*". Signal processing 86:33, 549-571, Elsevier Science, 2006.

Applications in geophysics

- manifold, e.g.
  - data sampling & regularization
    - Recovery of seismic data: practical considerations (today 11:15 am)
  - signal separation
    - Stable recovery and separation of seismic data (next talk)
    - Primary-multiple separation by curvelet frames (today 2:30 pm)
  - imaging
    - Compressed imaging (tomorrow 11:15 am)
- appealing
  - Stable Signal Recovery (SSR) theory provides **explicit recovery condition**.

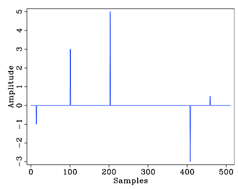
Agenda

- sparsity, compressibility & representations
- uncertainty principles and recovery
- stability

Sparsity

- what is sparsity?
  - the sparsity of a vector is defined as the number of 0 entries
    - sparse vector $\mathbf{x}$ (e.g. sparse spike train)
  - sparseness corresponds to small l_0 quasi-norm

    $\|\mathbf{x}\|_0 :=$ # of nonzero entries



- why using sparsity?
  - powerful property (i.e. extra piece of information about the signal) that offers striking benefits

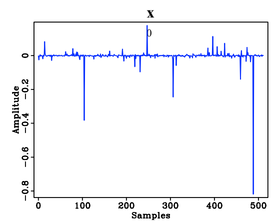
<u>note</u>:The idea of promoting sparsity for geophysical problems is commonly attributed to Claerbout and Muir in 1973. This was further developed e.g. by Oldenburg who proposed to deconvolve seismic traces for reflectivity as sparse spike trains. Although applications are successful in both cases, no explicit recovery condition is given.

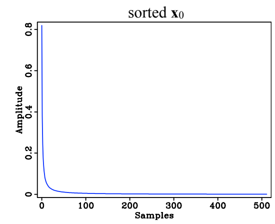
Compressibility

- what is compressibility?
  - $\mathbf{x}_0$ is compressible if its entries obey a power law

$$|\mathbf{x}_0|_{(k)} \leq C_r \cdot k^{-r}$$

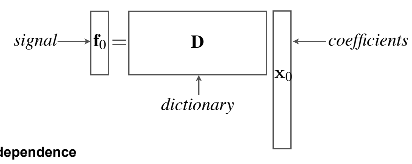
where $|\mathbf{x}_0|_{(k)}$ is the *k*-th largest value of $\mathbf{x}_0$ (e.g. wavelet coefficients of a piecewise smooth signal), C_r a constant, and $r \geq 1$.





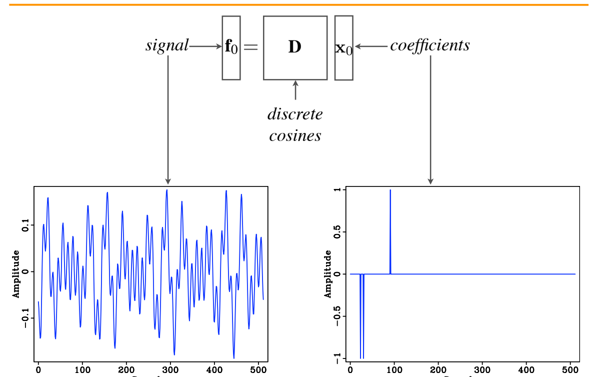
Sparse & compressible representations

- seek
  - **simplicity** (already forgot "Make it as simple as possible, but not simpler" by A. Einstein???)
    - signal $\mathbf{f}_0$ is built as a linear combination of **few atoms** from dictionary **D**

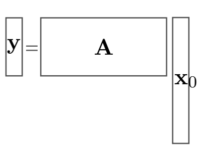


  - **independence**
    - atoms used to construct $\mathbf{f}_0$ do not contain redundant information
  - **expressiveness**
    - each selected atom significantly contributes to the construction of $\mathbf{f}_0$ (i.e. energy of the signal $\mathbf{f}_0$ is concentrated in **few significant coefficients**)

Example: sparse DCT representation



Perfect recovery

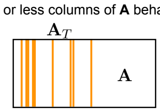


- conditions
  - **A** obeys a type of **uncertainty principle**
  - $\mathbf{x}_0$ is **sufficiently sparse**
- procedure

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y}$$

Closer look at the conditions

- uncertainty principles
  - Uniform Uncertainty Principle (UUP) is the most powerful
    - all subsets of *4S* or less columns of **A** behave like an **orthonormal basis**



- sparsity
  - $\mathbf{x}_0$ is an *S*-sparse vector (i.e. $\mathbf{x}_0$ has at most *S* nonzero entries)

<u>insight</u>: suppose

- $\mathbf{x}_0$ is an *S*-sparse vector
- **A** obeys the UUP for sets of size *2S*

compute $\mathbf{y} = \mathbf{A}\mathbf{x}_0$. Is $\mathbf{x}_0$ the sparsest solution that explains $\mathbf{y}$?

Consider another *S*-sparse vector $\mathbf{x}'_0$, then $\mathbf{h} = \mathbf{x}'_0 - \mathbf{x}_0$ have at most *2S* nonzeros and $\mathbf{A}\mathbf{h} = \mathbf{A}(\mathbf{x}'_0 - \mathbf{x}_0) = 0$. Since **A** obeys the UUP for sets of size *2S*, $\mathbf{A}_{|T|<2S}$ has no null space thus $\mathbf{x}'_0 = \mathbf{x}_0$.

Closer look at the recovery procedure

- severely underdetermined system of linear equations
  - infinitely many solutions
  - want the **sparsest** that explains the data ← RECOVERY PROCEDURE

$$\underbrace{\min_{\mathbf{x}} \|\mathbf{x}\|_1}_{\text{sparsity}} \quad \text{s.t.} \quad \underbrace{\mathbf{A}\mathbf{x} = \mathbf{y}}_{\text{perfect reconstruction}}$$

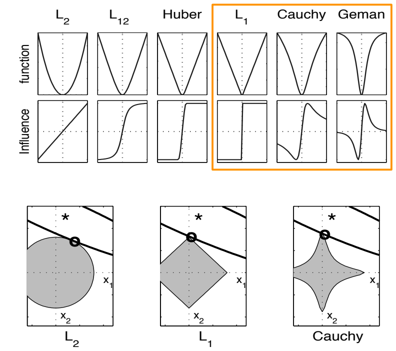
- performance
  - S-sparse vectors recovered from roughly on the order of S measurements (to within constant and *log* factors)

Note:

- sparsity measured using l_0 quasi-norm however recovery procedure is combinatorial (intractable!!)
- l_1 norm
  - is a **convex sparsity-enhancing** norm
  - is, under some circumstances, **equivalent to l_0 quasi-norm** for our recovery procedure
  - makes the **recovery procedure stable**

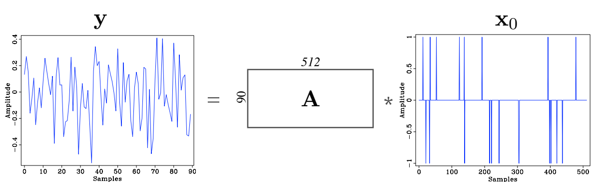
Figures courtesy P. Zwartjes

Sparsity-enhancing norms

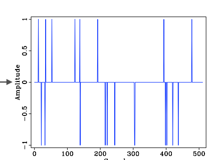


reproducible figure(s)

Example: perfect recovery



$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{y}$ →

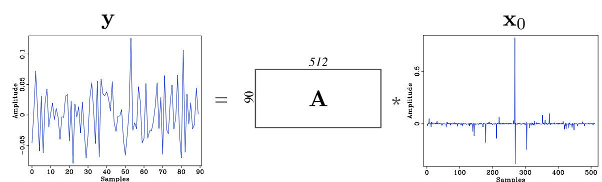


key facts:

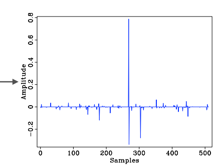
- $S = 20$
- ~ *90* observations required to perfectly recover
- more than 1:5 underdetermined

reproducible figure(s)

Example: recovery of comp. signals



$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{y}$ →

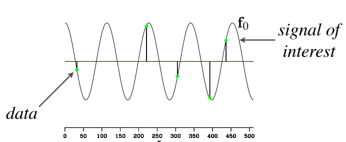


key facts:

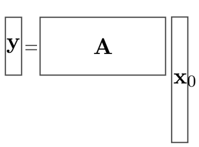
- $r = 1.1$
- ~ *90* observations
- more than 1:5 underdetermined

reproducible figure(s)

General setting



- $\mathbf{f}_0$ is not sparse/compressible but $\mathbf{f}_0$ has a sparse/compressible representation, i.e. $\mathbf{f}_0 = \mathbf{S}^H\,\mathbf{x}_0$.



with $\mathbf{A} := \mathbf{RMS}^H$ (*restriction*, *measurement*, *sparsity*)

General recovery condition

- performance

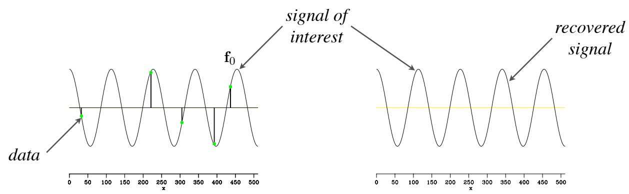
$$(\#\ \text{observations}) \propto \mu^2 \cdot (\#\ \text{nnz entries of } \mathbf{x}_0)$$

with μ the **mutual coherence** between **M** and **S**.

Note:

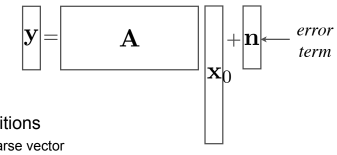
- mutual coherency measures the similarity of the columns of **M** and **S**.
- general recovery key factors:
  - mutual coherency
  - sparsity/compressibility of $\mathbf{x}_0$

Example: general perfect recovery



key facts:
- $S = 1$
- ~*5* observations required to perfectly recover
- more than 1:100 underdetermined

Stability

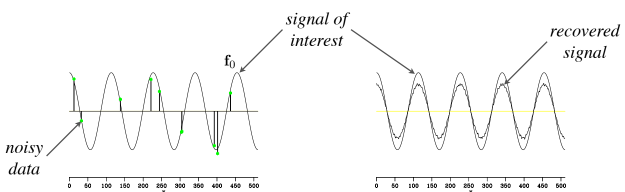


- recovery conditions
  - $\mathbf{x}_0$ is an S-sparse vector
  - **A** obeys the UUP for sets of size 4S
- procedure

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon$$

- performance
  - similar to noise-free but recovery error is on the order of the noise

Example: recovery from noisy data



key facts:
- $S = 1$
- ~*10* observations (SNR~8.3 dB)
- more than 1:50 underdetermined

Conclusions

- **sparsity** & **compressibility** are powerful properties (i.e. extra piece of information about the signal) that offer striking benefits
- stable signal recovery theory provides
  - **robust sampling criteria** as a function of
    - sparsity/compressibility
    - mutual coherence
  - a **stable** & **tractable recovery procedure** via l_1 minimization

Acknowledgments

- E. Candès and J. Romberg for l_1-magic
- S. Fomel and the developers of Madagascar

This overview work was carried out as part of the SINBAD project with financial support, secured through ITF, from the following organizations: BG, BP, Chevron, ExxonMobil, and Shell. SINBAD is part of a collaborative research & development (CRD) grant number 334810-05 funded by the Natural Science and Engineering Research Council (NSERC)