GRAVIMETRIC DETECTION BY COMPRESSED SENSING

Marina Meilă, Caren Marzban

University of Washington
Department of Statistics, Box 354322
Seattle WA 98195-4322

Ulvi Yurtsever

MathSense Analytics
1273 Sunny Oaks Circle
Altadena, CA 91001

ABSTRACT

We address the problem of identifying underground anomalies (e.g holes) based on gravity measurements.Our approach makes general assumptions about the shape of the hole, e.g that it can described by few wavelet coefficients. Such assumptions are known under the name of *sparsity assumptions*. Based on the recently developed compressed sensing (CS) methodology we output an estimate mass density over the whole domain of interest in one global optimization step. Our algorithms performance is promising on medium scale problems, even though the theoretical assumptions underlying CS do not hold for gravity problems of this kind.

Index Terms— Gravimetry, Compressed sensing

1. A PROBLEM: DISCOVERING UNDERGROUND FEATURES

This paper sets forth a novel approach for solving geospatial inversion problems using cutting edge techniques from statistics. Assume for instance that we want to obtain the mass density function in a given domain of interest from gravitational field measurements on the boundary of that domain. This is a well-studied and difficult problem. In all except a few special cases, the inverse problem has multiple solutions, and additional constraints (physical, or problem related) are needed to regularize it and to select a single, plausible solution [1].

In this paper we take a different slant on the problem: instead of seeking a complete description of the mass density $\rho(x, y, z)$ inside the domain of interest V, we will try to estimate only certain spatial features of the density. In point, we assume that discrete gravimetric measurements are taken on the ground, and we set out to map the near surface holes. We make one fundamental assumption: that the features to be detected, i.e. the holes, are described by a small number of parameters in the given representation. From an information p.o.v we seek an answer that can be represented with a small number of bits, assuming that one exists. In this case, we show experimentally that the problem has a unique solution, which can be obtained by a class of algorithms known as *compressed sensing* (CS).

Research supported by NURI grant HM1582-06-1-2035

The gravity inverse problem is the problem of finding the mass distribution ρ from measurements of the gravitational force G given in (1). Sometimes, only measurements of some components of the spatial gradient of G are available, as in (2).

$$G(r) \;=\; -G\int \frac{\rho(r')}{||r-r'||^2} d^3r' \tag{1}$$

$$(G_{xx} - G_{yy})(r) \;=\; 3G\int \frac{\rho(r')[(x-x')^2-(y-y')^2]}{||r-r'||^5} d^3r' \tag{2}$$

$$G_{xy}(r) \;=\; 3G\int \frac{\rho(r')(x-x')(y-y')}{||r-r'||^5} d^3r' \tag{3}$$

In practice, the unknown mass density is represented in some discretized form (e.g volume elements, 3D wavelets). For simplicity, assume that m_ξ is the mass of voxel $\xi = (x, y, z)$ of the domain V, containing a total of $|\Xi| = p$ unknown masses. The measurement vector is T with n elements (3 elements per location if T represents G and 2 if T represents the gradients of G). Because of superposition, the relation between the measurements and the masses is linear with a noise term: $T_j = \sum_{\xi\in\Xi} M_{j\xi} m_\xi + \tilde{\varepsilon}_j$ with $j = 1 : n$ and $\tilde{\varepsilon}_j$ the *noise*, accounting for measurement noise, discretization error, etc. In a realistic setting, p, *the number of mass elements*, can be larger than n, *the number of measurements*. Hence, the system of linear equations above is often underdetermined. Our task is to find the zeros of the mass vector. This problem is seemingly easier, but it still can have multiple solutions [2].

2. A COMPRESSED SENSING APPROACH

First we apply some simple preprocessing transformations to the data, which can be termed "substracting the background" (i.e the effect of any known mass on the measurement). The resulting system is

$$t \;=\; M\theta + \epsilon \tag{4}$$

where $t \in \mathbb{R}^n$ is the residual measurement, $\theta \in \mathbb{R}^p$ contains the unknown parameters (e.g $\theta_\xi = \bar{m} - m_\xi$), $M \in \mathbb{R}^{n\times p}$ is the measurement matrix, ϵ is the vector of residuals and $n < p$. Now our task is to find the *large components* of θ.

The compressed sensing theory of [3] refers to the *sparse (approximate) solutions* of (4). CS theory shows that when

the colums of M satisfy some near-orthogonality constraints, the noise ε_j is independent, and there is a sparse [1] solution θ^*, then the elements of θ^* above the noise level can be recovered by a simple linear program (LP).

Applied to our problem, the theorems above require thata the hole occupies relatively few of the volume elements, that the noise level is bounded. Both these conditions are often satisfied in real problems. The conditions referring to the measurement matrix M, however, are generally not satisfyied in real world settings [2].

Another requirement is that the noise values ϵ_j are independent. This requirement can be satisfied only if the ground density is a constant and θ represents the density, with no sparsifying transform. Otherwise, the variations of the ground density and the transform introduce error correlation.

The above considerations show that the theoretical results of [3] rarely, if ever, apply to our problem. In spite of this, we have modified and adapted the CS algorithm in a way that works surprisingly well in a variety of hole detection situations.

The algorithm requires only an estimate of $Var(\varepsilon)$ besides t and M and works for arbitrary hole shapes and positions. Note that the algorithm's goal is solely to map the hole, not to estimate the mass density. Hence, although the algorithm outputs a mass estimate for each voxel, the values that are not below the hole threshold should not be trusted. Figure 1,d shows an instance of the algorithm's output.

Our algorithm is described in Table 1.

Table 1. The HOLE DETECTION BY COMPRESSED SENSING Algorithm.

Given: $\phi \in \mathbb{R}^n$ measurement vector, $C \in \mathbb{R}^{n \times p}$ measurement matrix, $\bar{\rho}$ average mass and mass spatial covariance function

Preprocessing

1. Choose a sparsifying transform A so that $\rho = A\theta$ with θ sparse. (E.g A = identity, Haar wavelet)
2. $C \longleftarrow CA$ recompute measurement matrix
3. calculate variance σ of compound noise $\epsilon' \Rightarrow$ tolerance δ
4. Substract background $y \longleftarrow CA\mathbf{1}\bar{\rho} - \phi$

Sparse solution

5. Solve the compressed sensing LP obtain solution $\hat{\theta}$

Postprocessing

6. Reconstruct density $\rho = \bar{\rho} + A^{-1}\hat{\theta}$
7. Threshold and filter to detect the holes

3. EXPERIMENTS

Obtaining gravimetric field data *with ground truth* is expensive and difficult. Therefore, we present experiments on simulated data.

We artificially generate a measurement representing the gravity gradients $(G_{xx} - G_{yy}, G_{xy})$ at $75 = n/2$ locations at $z_0 = 0.5$ above a volume V of $p = 16 \times 16 \times 16 = 4096$ voxels. The volume contains either a Box hole or a Tunnel hole. The Box is a $h = 4 \times 2 \times 3 = 24$ voxel box, placed randomly at different locations in the volume. The (x, y, z) location of the box is considered to be that of the box center. The Tunnel is a tunnel parallel to the Y axis, with cross-section of 4×3 voxels; thus $h = 192$. The ground density is set at 1 with i.i.d noise uniformly distributed in $[-\sigma/2, \sigma/2]$. We ran experiments with $z_{\text{Box}}, z_{\text{Tunnel}} = 3, 4, 5$ and $\sigma = 0.1, 0.2\ 0.4$. Most ground densities have a relative range of variablility comparable to or lower than these values. We used the Haar wavelet transform on the data and for each of the Box and Tunnel and each parameter setting we generated 20 problem instances. A Boxand a Tunnel example are shown in Figure 1a, b.

Figure 1,d shows the algorithm's output for a Box. The blue line is the estimated mass for each voxel[3]. This is thresholded and values lower than the threshold are marked as "hole". These are shown in green. The true hole voxels are marked with red on the brown curve representing the true mass. Note that imposing sparsity on the result makes the mass estimate be 0 at most locations. One can also see that with a different threshold, many of the current *False Positives (FP)* could have been corrected. Therefore, we are currently working on an improved thresholding algorithm.

Figure 1,c shows the spatial reconstruction of the Box, with the hole voxels in gray. The true hole is completely covered, but there are many FP's outside. This is partly an effect of the wavelet transform, which will mark whole regions as "hole" and "full" rather than individual voxels. One can also notice that the X, Y localization of the hole, computed as the X, Y mean of the hole voxels, is relatively good, but the Z localization is poor. This is typical of all our experiments.

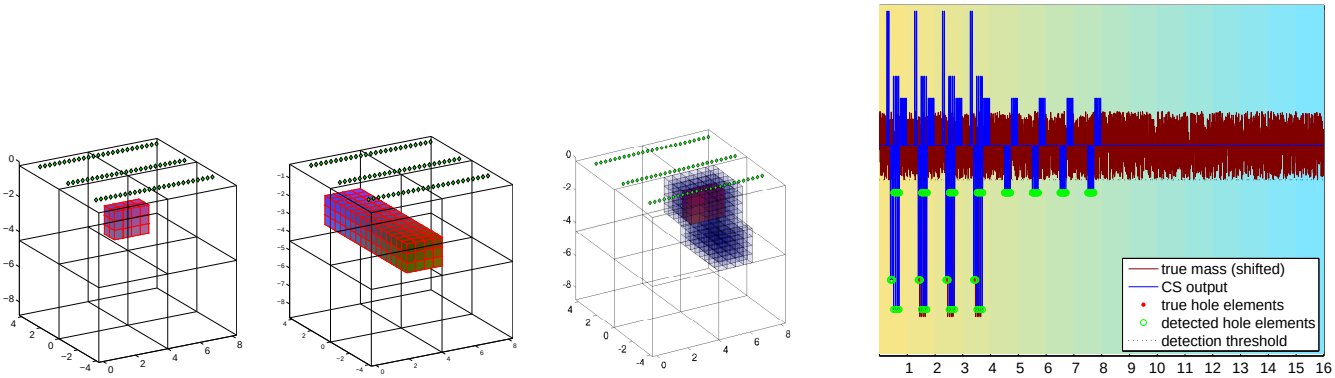
Figure 2a,c shows X for an example localization errors for Box and Tunnel on all 20 runs for $z = 4$, $\sigma = 0.2$. In all but 1 cases the error is within 1 unit from the truth. Results have little sensitivity to σ, but they degrade with depth, and a hole under $z = 6$ can rarely be detected. The results are more accurate for Tunnel, suggesting that a hole with more volume may be more detectable. Figure 2b, d show voxel level error rates. The error rate is defined w.r.t the **hole volume**

$$FN = \frac{\text{\# undetected hole voxels}}{\text{\# hole voxels}} \leq 1 \qquad (5)$$

[1] Sparsity of θ is defined alternatively by small l-norm, for some $l \leq 1$ or by θ having few non-zero entries.

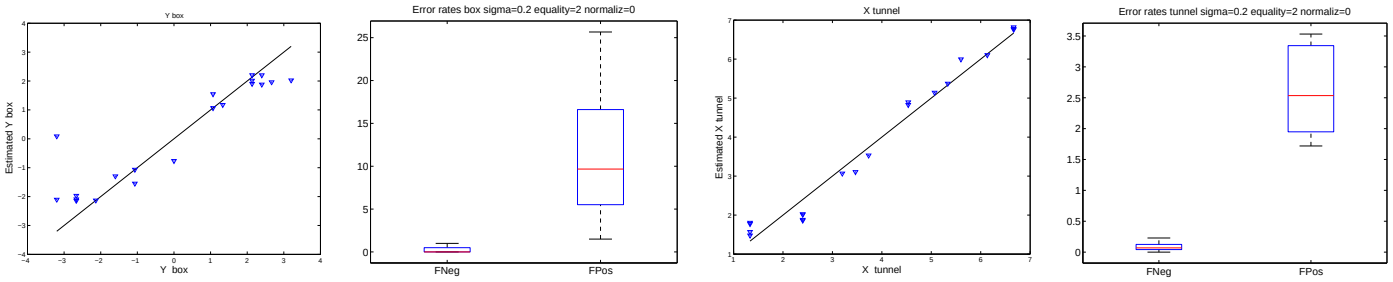
[2] This can be seen by noting that two adjacent columns of M can be very close to each other

[3] Because of substracting the background the average mass estimate is shifted to 0.



a Box example **b** Tunnel example **c** Box reconstruction (grey) superimposed on truth **d** HDCS output superimposed on truth

Fig. 1. Examples of Box **(a)** and Tunnel **(b)**, a Box reconstruction and true hole superimposed **c** and the result of the HDCS Algorithm compared to the truth: true mass (brown), reconstructed mass (blue), true hole voxels (red), and estimated hole voxels (green).



Box X location Box error rate Tunnel X location Box error rate

Fig. 2. X Localization, false positive (FP) and false negative (FN) rates for a box hole of size $4 \times 2 \times 3$ (left) and a tunnel hole of size 4×3 units (right). The holes are placed in random (x, y) positions, at depth $z = 4$ from the surface; noise level $\sigma = 0.2$.

Table 2. Hole detection probabilities for a toy example (left) and the Box described in text (right) estimated from 100 random runs, 50 containing a box and 50 containing no hole. The errors are estimated from the valid runs only.

Toy example

σ	False alarms %	Misses %	Valid runs %
0.05	0.12	0	84
0.10	0.11	0	89
0.20	0.08	0	85
0.30	0.10	0	88
0.40	0.13	0	84
0.80	0.16	0	81

Large example

$\sigma = 0.4$, $h = 24$, $n = 234$, $p = 4096$

depth z	False alarms %	Misses %	Valid runs%
3	0.08	0	95
4	0.03	0	83
5	0.05	0.17	86

$$FP = \frac{\text{\# falsely detected voxels}}{\text{\# hole voxels}} \quad (6)$$

hence the FP rate can be larger than 1. The FN rate is very good, while the FP rate is relatively large.

We also ran a set of detection experiments. In these, 50 problem instances containing a Box in a random position and 50 instances containing no hole were generated. The results are shown in Table 2 along with results from a similar but much smaller experiment, having $p = 180$, $h = 24$, $n = 24$ and random z. Some runs did not terminate with convergence of the LP solver, and those were labeled *invalid*. The error rates, *False alarm*, i.e detection of an inexisting hole and failing to detect, i.e *Missing* and existing hole, were computed only on the *valid* runs. One again notes the robustness to the noise level (in the toy case). Therefore, in the second case, we ran the experiments only for the largest σ, 0.4. The results also seem robust to the size of the problem, but there are signs of degradation with the depth z.

4. DISCUSSION

The approach presented here is extremely versatile. It can be immediately adapted to other types of gravimetric instruments, other types of fields (e.g magnetic), and to other features than holes. In fact, our algorithm can robustly detect any feature that can be represented by a superposition of matched linear filters.

In the current experiments, we have made *no physical assumptions about the data* except for sparsity. This is one of the great strengths of compressed sensing. However, it is possible to further improve performance by incorporating physical knowledge, and this is our current objective.

We are not aware of any solution to hole detection that imposes *no shape constraints* on the hole. The vast majority of CS applications rely on a user-designed measurement matrix M (this is the case in data compression), on very large values of p (at least 10^3), and on much larger n/h ratios (here h is the number of non-zero parameters, i.e the number of hole voxels). In contrast, our M is given by the geometry of the problem, and the experiments are pushing the envelope of CS algorithms far beyond anything we have encountered (we successfuly mapped "tunnels" with $n/h = 24/30 < 1$). This is a powerful new tool for a variety of detection problems.

5. REFERENCES

[1] Victor Isakov, *Inverse problems for partial differential equations*, Number 127 in Applied Mathematical Science. Springer, second edition, 2006.

[2] Victor Isakov, *Inverse source problems*, Number 34 in Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 1990, QA 374.I83 1990.

[3] David Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.