Manifold Denoising by Nonlinear Robust Principal Component Analysis

He Lyu, Ningyu Sha, Shuyang Qin, Ming Yan, Yuying Xie, Rongrong Wang

Michigan State University

CMSE

Summary and Contributions

This work concerns the problem of extending Robust PCA to manifold setting, where the observed data is the sum of a sparse component and a component drawn from some low dimensional manifold.

We propose an optimization framework that separates the sparse component from the manifold under noisy data.

- A theoretical guarantee for the method
- A curvature estimation method that may be of independent interest

Low Frequency Asymptotics

By Lemma 3.1 in [1], it has been shown that:

The **unique** solution $u(x,k) \in H^1_{loc}(\mathbb{R}^3)$ to the Helmholtz equation is given by the integral form:

$$u(x,k) = -k^2 \int_{\mathbb{R}^3} [1 - c^{-2}(y)] u(y,k) \Phi(x-y) dy - \frac{ik}{2\pi} \int_{\mathbb{R}^3} f(y) c^{-2}(y) \Phi(x-y) dy.$$

Moreover, as $\Phi(x)$ allows the following asymptotic expansion:

$$\Phi(x) = \Phi_0(x) + \mathcal{O}(k), \quad \text{as } k \to 0^+,$$

the integral solution also has an asymptotic form:

$$u(x,k) = -\frac{ik}{2\pi} \int_\Omega \frac{f(y)}{c^2(y)} \Phi_0(x-y) dy + \mathcal{O}(k^2).$$

Furthermore, in [2], giving 2 groups of parameters sharing the same boundary measurement,

$$\int_\Omega \left(\frac{f(x)}{c^2(x)} - \frac{\tilde{f}(x)}{\tilde{c}^2(x)} \right) \varphi(x) dx = 0,$$
$$\int_\Omega \left(\frac{1}{c^2(x)} - \frac{1}{\tilde{c}^2(x)} \right) \varphi(x) dx = 0.$$

The uniqueness of both point source and low frequency wave speed can be guaranteed by above.

Proposed Method

Based on the theoretical analysis as well as Green's Theorem, as $k \to 0^+$:

$$-\frac{ik}{2\pi} \int_\Omega \frac{f(x)}{c^2(x)} \varphi(x) dx = \int_{\partial\Omega} u(x,k) \partial_\nu \varphi(x) d\sigma - \int_{\partial\Omega} \partial_\nu u(x,k) \varphi(x) d\sigma,$$

where the R.H.S. can be calculated by measurement data. For simplicity, we construct harmonic sensing matrix by choosing random shifts to the complex harmonic polynomial of order 1, i.e.:

$$\varphi_{(x_0,y_0)}(x,y) = (x_0 - x) + i(y_0 - y),$$

and then evaluate such functions at discretized grids. Here let shifts $(x_0, y_0) \sim \text{Unif}([-1,1]^2)$, uniformly sampling from computational domain.

Low Frequency Wavespeed

The same method can be applied to reconstruct wavespeed. The speed itself can not be sparse, but it may have a sparse representation under Fourier basis.

Figure 1: An example of low frequency wavespeed

Numerical Experiments

50 shifted harmonic functions, measurement data calculated by integration with 5% Gaussian noise:

Figure 2: Recovered wavespeed with first 2 frequencies

Figure 3: Recovered wavespeed with first 4 frequencies

Figure 4: Source Location Reconstruction

	1	2	3	4
true	0.615	0.577	0.625	0.447
reconstructed	0.526	0.440	0.469	0.321

Table 1: Reconstructed and true source intensities

20 shifted harmonic functions, 5% Gaussian noise and reconstructed medium in Figure 3.

The error of estimating source intensity comes mostly from reconstruction error of low frequency wave speed.

Incoherence Conditions

By taking random shifts for each row, and normalizing columns, the incoherence condition for 1-sparse recovery holds.

Theorem 1 (1-Incoherence Condition) *By our way of constructing sensing matrix, with probability at least $1 - 3\exp(-cm)$, where m is the number of shifts to be chosen, the coherence between different columns satisfy:*

$$\max_{i \neq j} |\mu_{i,j}| \le 1 - \frac{1}{32} h^2.$$

Thus for 1-sparse source term, the location of the point source can be reconstructed with probability at least $1 - \delta$ by $c \log \delta$ harmonic functions.

Conclusion & Future Work

- We demonstrated the ability of sparse recovery using harmonic function, numerically tested the performance of reconstructing point sources. The result is robust w.r.t small noise.
- Choose appropriate forward modeling methods to measure boundary data with singularity in initial source, and factor out the part of data corresponding to those sparse terms.
- Choice of harmonic functions that have better property for sparse reconstruction.
- Only partial data is measured on the boundary, choose appropriate harmonic functions to account for information loss.

References

[1] Hongyu Liu and Gunther Uhlmann. Determining both sound speed and internal source in thermo- and photo-acoustic tomography. Inverse Problems, 2015.

[2] Christina Knox and Amir Moradifam. Determining both the source of a wave and its speed in a medium from boundary measurements. arXiv preprint, arXiv:1803.06750.