

Signal Recovery on a Manifold

Yang Wang

The Hong Kong University of Science and Technology

Joint work with Yi Rong and Zhiqiang Xu

May 19, 2017



1 The Problem

- Stating the Problem
- Examples
- Injectivity and Minimality Questions

2 Connection to Algebraic Geometry

- Basics
- Dimension and Intersection

3 Putting Things Together for the Signal Recovery Problem

- Ideas Through a Simple Example
- Full Recovery on Variety from Random Measurements
- Admissibility Condition
- Main Theorems

Stating the Problem

Solving for a system of linear equations is perhaps the most fundamental mathematical problem in just about all areas of science and engineering applications. Here we have an unknown vector $\mathbf{x} \in \mathbb{F}^d$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. We would like to recover $\mathbf{x}$ through a set of linear measurements $A\mathbf{x}$. In other words,

$$\text{Solve: } A\mathbf{x} = \mathbf{b}.$$

The problem is that there are some constraints:

- $A \in \mathbb{F}^{N \times d}$ where N may be significantly smaller than d .
- $\mathbf{x}$ lies on some nonlinear manifold $\mathcal{M}$.

Example 1: Compressive Sensing

This is probably the example that is most familiar to people. It arises in numerous applications. The setup is that we would like to solve for $A\mathbf{x} = \mathbf{b}$ where $\mathbf{x}$ is *sparse*.

- $A \in \mathbb{F}^{N \times d}$ where N maybe significantly smaller than d .
- $\mathbf{x} \in \mathbb{F}_k^d$, the set of all vectors in $\mathbb{F}^d$ with sparsity $0 \leq k \ll d$.

Note that $\mathbb{F}_k^d$ is nonlinear.

Example 2: Low Rank Matrix Recovery

This problem has gained a lot of attention in recent years due to many applications in various fields like online recommendation systems, image processing, etc. Here our unknown “vector” is a matrix $X = (x_{mn}) \in \mathbb{F}^{p \times q}$ with $\text{rank}(X) \leq r$. We are given a set of linear measurements $\{L_j(X) = b_j\}_{j=1}^N$. We wish to recover X .

- If $L_j(X) = x_{m_j n_j}$ then we are looking at the Netflix Problem of recovering missing entries.
- More generally each $L_j(X) = \text{tr}(A_j^T X)$ for some $A_j \in \mathbb{F}^{p \times q}$. Every linear function on $\mathbb{F}^{p \times q}$ can be expressed in this form. For example, $L(X) = x_{mn} = \text{tr}(E_{mn}^T X)$.

Note here the manifold is $\mathcal{M} = \{X \in \mathbb{F}^{p \times q} : \text{rank}(X) \leq r\}$.

Example 3: Phase Retrieval

This problem has been extensively studied due to its numerous applications, such as X-ray diffraction and other imaging applications, communication, waveform, etc. Here we would like to recover an unknown $\mathbf{x} \in \mathbb{F}^d$ through a set of *quadratic measurements* $b_1 = |\langle \mathbf{f}_1, \mathbf{x} \rangle|^2, b_2 = |\langle \mathbf{f}_2, \mathbf{x} \rangle|^2, \dots, b_N = |\langle \mathbf{f}_N, \mathbf{x} \rangle|^2$.

“Linearization”: Set $X = \mathbf{x}\mathbf{x}^*$ and $A_j = \mathbf{f}_j\mathbf{f}_j^*$. Then we have

$$\begin{aligned} b_j &= |\langle \mathbf{f}_j, \mathbf{x} \rangle|^2 = \mathbf{x}^* \mathbf{f}_j \mathbf{f}_j^* \mathbf{x} \\ &= \mathbf{x}^* A_j \mathbf{x} = \text{tr}(A_j \mathbf{x} \mathbf{x}^*) = \text{tr}(A_j X). \end{aligned}$$

So we are now recovering $X \in \mathbb{F}^{d \times d}$ from a set of linear measurements, with X lying on the manifold

$$\mathcal{M} = \{X \in \mathbb{F}^{d \times d} : X \geq 0, \text{rank}(X) \leq 1\}.$$

Example 4: Other Matrix Retrieval Problem

An example of this type of problems is the Projection Retrieval problem, where we would like to recover an unknown orthogonal projection matrix $P \in \mathbb{F}^{d \times d}$ of rank r through a set of measurements via sample points $\mathbf{x}_j \in \mathbb{F}^d$:

$$b_1 = \|P\mathbf{x}_1\|^2, b_2 = \|P\mathbf{x}_2\|^2, \dots, b_N = \|P\mathbf{x}_N\|^2.$$

Again, set $A_j = \mathbf{x}_j \mathbf{x}_j^*$. Then

$$b_j = \|P\mathbf{x}_j\|^2 = \mathbf{x}_j^* P P^* \mathbf{x}_j = \mathbf{x}_j^* P \mathbf{x}_j = \text{tr}(\mathbf{x}_j \mathbf{x}_j^* P) = \text{tr}(A_j P).$$

So we are recovering $P \in \mathbb{F}^{d \times d}$ from a set of linear measurements, with P lying on the manifold $\mathcal{M}$ of **all orthogonal projection matrices of rank at most r** . Of course, here one can **replace $\mathcal{M}$ by other manifolds**.

Example 5: The Missing Distance Problem

Let $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{F}^d$ be a finite set of points and set $d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|$. It is well-known that the point cloud $(\mathbf{x}_j)$ is uniquely determined up to an isometry by these distances (d_{ij}) . But *what if some d_{ij} are missing?* Without loss of generality let $\mathbf{x}_0 = 0$. Set

$$X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{F}^{d \times N}, \quad P := X^* X = [\langle \mathbf{x}_i, \mathbf{x}_j \rangle]$$

Note that each d_{ij} is a **linear function** of P :

$$d_{ij}^2 = \|\mathbf{x}_i - \mathbf{x}_j\|^2 = \langle \mathbf{x}_i, \mathbf{x}_i \rangle + \langle \mathbf{x}_j, \mathbf{x}_j \rangle - 2\langle \mathbf{x}_i, \mathbf{x}_j \rangle.$$

Thus the Missing Distance problem is equivalent to recovering $P \in \mathbb{F}^{d \times d}$ from a set of linear measurements, with P lying on the manifold

$$\mathcal{M} = \{P \in \mathbb{F}^{n \times n} : P \geq 0, \text{rank}(P) \leq d\}.$$

Injectivity

Let $L_j(\mathbf{x}), j = 1, \dots, N$ be linear functions on $\mathbb{F}^d$. Define $\mathbf{L} : \mathbb{F}^d \longrightarrow \mathbb{F}^N$

$$\mathbf{L}(\mathbf{x}) := (L_1(\mathbf{x}), L_2(\mathbf{x}), \dots, L_N(\mathbf{x}))^T.$$

Definition

Let $\mathcal{M} \subseteq \mathbb{F}^d$. We say $\{L_j(\mathbf{x})\}_{j=1}^N$ has the $\mathcal{M}$ -recovery property if the map $\mathbf{L}$ is injective on $\mathcal{M}$. We say it has the almost everywhere $\mathcal{M}$ -recovery property if $\mathbf{L}^{-1}(\mathbf{L}(\mathbf{x})) = \{\mathbf{x}\}$ for almost all $\mathbf{x} \in \mathcal{M}$.

- $\mathcal{M}$ -recovery property $\Leftrightarrow$ every $\mathbf{x} \in \mathbb{F}^d$ can be recovered from $\mathbf{L}(\mathbf{x})$.
- Almost everywhere $\mathcal{M}$ -recovery property $\Leftrightarrow$ almost all $\mathbf{x} \in \mathbb{F}^d$ can be recovered from $\mathbf{L}(\mathbf{x})$.

Minimality Questions

- What is the minimal number N for which a set of linear measurements $\{L_j(\mathbf{x})\}_{j=1}^N$ can have the $\mathcal{M}$ -recovery property?
- If we randomly pick our linear measurements $\{L_j(\mathbf{x})\}_{j=1}^N$, how many are needed for them to have the $\mathcal{M}$ -recovery property with “high” probability?
- What is the minimal number N for which a set of linear measurements $\{L_j(\mathbf{x})\}_{j=1}^N$ can have the almost everywhere $\mathcal{M}$ -recovery property?
- If we randomly pick our linear measurements $\{L_j(\mathbf{x})\}_{j=1}^N$, how many are needed for them to have the $\mathcal{M}$ -recovery property with “high” probability?

These questions for some of the aforementioned examples are answered, but many remain unresolved.

Background from Algebraic Geometry

It turns out that it is more appropriate to discuss projective varieties rather than manifolds. Here we shall define a *projective variety* in $\mathbb{C}^d$ to be the zero locus of a finite set of homogeneous polynomials.

- The set of all *symmetric matrices* in $\mathbb{C}^{p \times p}$ is a projective variety, which can be defined by homogeneous polynomials $A = A^T$. However, the set of all *Hermitian matrices* in $\mathbb{C}^{p \times p}$ is not.
- The set of all rank r or less matrices in $\mathbb{C}^{p \times q}$ is a projective variety, called the *determinantal variety*. It has dimension $(p + q)r - r^2$.
- The set of all rank r or less matrices in $\mathbb{C}^{p \times q}$ satisfying $A^T = A$ and $A^2 = cA$ for some $c \in \mathbb{C}$ is a projective variety.

Basic Concepts: Irreducibility, dimension, union and intersection.

Dimension and Intersection

Proposition

Let $\mathcal{M}$ be a projective variety and $\mathcal{P}$ be a hyperplane in $\mathbb{C}^d$. Then $\dim(\mathcal{M} \cap \mathcal{P}) \geq \dim(\mathcal{M}) - 1$. Furthermore, if $\mathcal{P}$ does not contain an irreducible component of $\mathcal{M}$ then $\dim(\mathcal{M} \cap \mathcal{P}) = \dim(\mathcal{M}) - 1$.

Thus a hyperplane $\mathcal{P}$ in “generic position” will have $\dim(\mathcal{M} \cap \mathcal{P}) = \dim(\mathcal{M}) - 1$. Note the result doesn't hold for real projective varieties.

Definition

*Let $\mathcal{M}$ be a projective variety in $\mathbb{C}^d$ with $\dim \mathcal{M} > 0$ and let $\{\ell_\alpha(\mathbf{x}) : \alpha \in I\}$ be a family of (homogeneous) linear functions. We say $\mathcal{M}$ is **admissible** with respect to $\{\ell_\alpha(\mathbf{x})\}$ if $\dim(\mathcal{M} \cap \{\ell_\alpha(\mathbf{x}) = 0\}) < \dim \mathcal{M}$ for all $\alpha \in I$.*

Dimension and Intersection

Fortunately the admissibility condition is easy to verify and satisfy. In just about all cases we are interested in studying, the condition holds.

Another extremely useful result, which we need to use to handle recovery problems on $\mathbb{R}^d$, is the following:

Proposition

Let $\mathcal{M}$ be a projective variety in $\mathbb{C}^d$. Let $\mathcal{M}_{\mathbb{R}}$ be the restriction of $\mathcal{M}$ on the reals, i.e. $\mathcal{M}_{\mathbb{R}} = \mathcal{M} \cap \mathbb{R}^d$. Then $\mathcal{M}_{\mathbb{R}}$ is a real projective variety and its real dimension satisfies $\dim_{\mathbb{R}}(\mathcal{M}_{\mathbb{R}}) \leq \dim(\mathcal{M})$.

We can also take a real projective variety V in $\mathbb{R}^d$ and “lift” it to a projective variety $\bar{V}$ in $\mathbb{C}^d$. Clearly we have $\dim(\bar{V}) \geq \dim_{\mathbb{R}}(V)$.

Signal Recovery: Ideas from an Example

Let $\mathcal{M} = \mathcal{M}_{p \times q, r}(\mathbb{C})$, the set of all matrices in $\mathbb{C}^{p \times q}$ with rank $r < \min(p, q)/2$.

Let $A_1, \dots, A_N$ be i.i.d. Gaussian random matrices in $\mathbb{C}^{p \times q}$ and for $X \in \mathcal{M}$ consider the linear measurements

$$\mathbf{L}(X) := \left(\text{tr}(A_1^T X), \text{tr}(A_2^T X), \dots, \text{tr}(A_N^T X) \right).$$

How big should N be for $\mathbf{L}$ to be injective on $\mathcal{M}$?

This is equivalent to $\mathbf{L}(X - Y) = 0$ iff $X - Y = 0$ where $X, Y \in \mathcal{M}$, or in other words, $\mathbf{L}(Z) = 0$ iff $Z = 0$ where $Z \in \mathcal{M} - \mathcal{M} = \mathcal{M}_{p \times q, 2r}(\mathbb{C})$.

Theorem (Xu 2015)

$\mathbf{L}$ is injective $\Rightarrow N \geq 2r(p + q) - 4r^2 = \dim(\mathcal{M} - \mathcal{M})$.

Signal Recovery: Ideas from an Example

Also proved in [Xu 2015]

- (1) The result is sharp. $N \geq 2r(p + q) - 4r^2$ random measurements give $\mathcal{M}$ -recovery property with probability 1.
- (2) The result is false if we change $\mathbb{C}$ to $\mathbb{R}$.

Reasons: (1) Each random measurement defines a hyperplane in generic position in the projective space. So its intersection with a given projective variety cuts the dimension by exactly one. After $N = 2r(p + q) - 4r^2$ intersections the dimension becomes 0.

(2) The dimension drop for the intersection can be > 1 for real projective variety. Thus result is false in the real case. Counter example was given.

$\mathcal{M}$ -Recovery from random measurements

Theorem (YW and Xu 2016)

Let $\mathcal{M}$ be a projective variety in $\mathbb{C}^{p \times q}$ with $\dim(\mathcal{M}) = K$. Let $A_1, A_2, \dots, A_N$ be randomly chosen matrices in $\mathbb{C}^{p \times q}$ according to some absolutely continuous probability distribution. Set

$$\mathbf{L}(X) := \left(\text{tr}(A_1^T X), \text{tr}(A_2^T X), \dots, \text{tr}(A_N^T X) \right).$$

- (1) If $N < K$ then $\mathbf{L}$ is not injective on $\mathcal{M}$. Thus $\{\text{tr}(A_i^T X)\}_{i=1}^N$ does not have the $\mathcal{M}$ -recovery property.
- (2) If $N \geq K$ then $\mathbf{L}$ is injective on $\mathcal{M}$ with probability 1. Thus $\{\text{tr}(A_i^T X)\}_{i=1}^N$ has the $\mathcal{M}$ -recovery property with probability 1.

Diving Deeper into the Example

In the previous example of recovering rank r matrices, what if we restrict the measurement matrices to some special manifold (projective variety)?

For example, suppose we require $A_1, A_2, \dots, A_N$ be randomly chosen rank 1 matrices in $\mathbb{C}^{p \times q}$, will $\mathbf{L}(X)$ still be injective?

The answer is still **Yes!** Note here A_j lie on a lower dimensional variety so we can no longer use the *generic position* argument. We need to have a “generic position” argument on a lower dimensional variety. This is done through the **Admissibility** condition.

Admissibility Condition

Definition (YW and Xu 16)

Let V be a projective variety in $\mathbb{C}^d$ with $\dim V > 0$ and let $\{\ell_\alpha(\mathbf{x}) : \alpha \in I\}$ be a family of linear functions on $\mathbb{C}^d$. We say V is *admissible* with respect to $\{\ell_\alpha(\mathbf{x})\}$ if $\dim(V \cap \{\ell_\alpha(\mathbf{x}) = 0\}) < \dim V$ for all $\alpha \in I$.

Without getting into the technical details, the basic message is that the admissibility condition can be rather easily checked, and it is satisfied in virtually all cases we care about.

There are definitely counterexamples, e.g. V is the set of all symmetric matrices and ℓ_α is of the form $\text{tr}(A^T X)$ where A is skew-symmetric.

General Setup of the Problem

Let $\mathcal{M} \subseteq \mathcal{F}^{p \times q}$ be a projective variety. Let $A_1, A_2, \dots, A_N \in \mathbb{F}^{p \times q}$ and set

$$\mathbf{L}(X) := \left(\text{tr}(A_1^T X), \text{tr}(A_2^T X), \dots, \text{tr}(A_N^T X) \right).$$

Assume that each A_j is restricted to some subset V_j of $\mathcal{F}^{p \times q}$. When will $\mathbf{L}$ be injective on $\mathcal{M}$, namely, when will the $\mathcal{M}$ -recovery property hold? And when will the almost everywhere $\mathcal{M}$ -recovery property hold?

Through proper transformations we can always assume the above setup without loss of generality.

Main Result

Theorem (YW and Xu 16)

Let $\mathcal{M}$ be a projective varieties in $\mathbb{C}^{p \times q}$. Let $A_j \in V_j$ be generic, where V_j is a projective variety satisfying “appropriate” admissibility condition. Then for $N \geq \dim(\mathcal{M} - \mathcal{M})$

$$\mathbf{L}(X) := \left(\text{tr}(A_1^T X), \text{tr}(A_2^T X), \dots, \text{tr}(A_N^T X) \right)$$

is *injective on $\mathcal{M}$* . In other words, the maps $\{\text{tr}(A_j^T X)\}_{j=1}^N$ have the $\mathcal{M}$ -recovery property.

On the other hand, if $N < \dim(\mathcal{M} - \mathcal{M})$ then $\mathbf{L}$ is not injective on $\mathcal{M}$.

- Here “appropriate” admissibility condition is in fact quite explicit and easily checked. We skip the details.

Main Result: Real Case

Theorem (YW and Xu 16)

Let $\mathcal{M}$ be a projective varieties in $\mathbb{R}^{p \times q}$. Let $A_j \in V_j$ be generic, where V_j is a projective variety in $\mathbb{R}^{p \times q}$. Then for $N \geq \dim(\mathcal{M} - \mathcal{M})$

$$\mathbf{L}(X) := \left(\text{tr}(A_1^T X), \text{tr}(A_2^T X), \dots, \text{tr}(A_N^T X) \right)$$

is *injective on $\mathcal{M}$* , provided that $\bar{\mathcal{M}}$ and $\bar{V}_j$ satisfy “appropriate” admissibility condition and

$$\dim_{\mathbb{R}}(\mathcal{M} - \mathcal{M}) = \dim(\bar{\mathcal{M}} - \bar{\mathcal{M}}), \quad \dim_{\mathbb{R}}(V_j) = \dim(\bar{V}_j).$$

- The additional conditions are typically satisfied by the cases we care about.

Application to Various Examples

- For $\mathcal{M}$ = the set of all rank r matrices in $\mathbb{F}^{p \times q}$, $r \leq \min(p, q)/2$, we have $N \geq 2(p + q)r - 4r^2$. For complex this is sharp, but for real it is not.
- For $\mathcal{M}$ = the set of all rank r symmetric matrices in $\mathbb{F}^{p \times p}$, $r \leq p/2$, we have $N \geq r(2p - 2r + 1)$. For the set of $p \times p$ Hermitian matrices, we have $N \geq 4r(p - r)$ (not sharp).
- For compressive sensing, we have $N \geq 2k$ where k is the sparsity (sharp).
- For phase retrieval in $\mathbb{F}^d$ we have $N \geq 2d - 1$ for $\mathbb{F} = \mathbb{R}$ (sharp) and $N \geq 4d - 4$ for $\mathbb{F} = \mathbb{C}$ (not sharp).
- There are many other examples that similar result can be established.

Lower Bound

In the real case the results from dimension and intersection theory from algebraic geometry are usually not sharp. This leads to several open questions:

- Can we find sharp lower bound for N ? This appears to be intractable in general, but in many special cases sharp lower bound can be obtained.
- Lower bound can be obtained using the results on manifold embedding into Euclidean space. This is a classic area and is difficult for people not in the field. But we can apply known results. Typically we get a gap of $O(\log d)$.
- It is also conjectured that generically the N we obtain through this framework is optimal.
- I shall skip the details on the lower bound question.

Almost Everywhere $\mathcal{M}$ -Recovery: Complex Case

Theorem (Rong, YW and Xu 17)

Let $\mathcal{M}$ be a projective varieties in $\mathbb{C}^{p \times q}$. Let $A_j \in V_j$ be generic, where V_j is a projective variety satisfying “appropriate” admissibility condition. Then for $N > \dim(\mathcal{M})$

$$\mathbf{L}(X) := \left(\text{tr}(A_1^T X), \text{tr}(A_2^T X), \dots, \text{tr}(A_N^T X) \right)$$

is *almost everywhere injective on $\mathcal{M}$* . In other words, the maps $\{\text{tr}(A_j^T X)\}_{j=1}^N$ have the almost everywhere $\mathcal{M}$ -recovery property.

On the other hand, if $N < \dim(\mathcal{M})$ then $\mathbf{L}$ is not injective on $\mathcal{M}$.

- Here “appropriate” admissibility condition is the same as in the previous theorem

Almost Everywhere $\mathcal{M}$ -Recovery: Real Case

Theorem (Rong, YW and Xu 17)

Let $\mathcal{M}$ be a projective varieties in $\mathbb{R}^{p \times q}$. Let $A_j \in V_j$ be generic, where V_j is a projective variety in $\mathbb{R}^{p \times q}$. Then for $N > \dim(\mathcal{M})$

$$\mathbf{L}(X) := \left(\text{tr}(A_1^T X), \text{tr}(A_2^T X), \dots, \text{tr}(A_N^T X) \right)$$

is *almost everywhere injective on $\mathcal{M}$* , provided that $\bar{\mathcal{M}}$ and $\bar{V}_j$ satisfy “appropriate” admissibility condition and

$$\dim_{\mathbb{R}}(\mathcal{M} - \mathcal{M}) = \dim(\bar{\mathcal{M}} - \bar{\mathcal{M}}), \quad \dim_{\mathbb{R}}(V_j) = \dim(\bar{V}_j).$$

On the other hand, if $N < \dim(\mathcal{M})$ then $\mathbf{L}$ is not injective on $\mathcal{M}$.

- The additional conditions are the same as in previous theorem.

Various Examples of Almost Everywhere $\mathcal{M}$ -Recovery

- For $\mathcal{M}$ = the set of all rank r matrices in $\mathbb{F}^{p \times q}$, $r \leq \min(p, q)/2$, we have $N > (p + q)r - r^2$.
- For $\mathcal{M}$ = the set of all rank r symmetric matrices in $\mathbb{F}^{p \times p}$, $r < p$, we have $N > r(2p - r + 1)/2$. For the set of $p \times p$ Hermitian matrices, we have $N > r(2p - r)$.
- For compressive sensing, we have $N > k$ where k is the sparsity.
- For phase retrieval in $\mathbb{F}^d$ we have $N > d$ for $\mathbb{F} = \mathbb{R}$ (sharp) and $N \geq 2d$ for $\mathbb{F} = \mathbb{C}$.
- Not known for many of the above whether they are sharp.

How About Computation?

- These results do not address the computational side of things. Generally it is a challenging problem.
- For many problems it is feasible computationally to perform recovery where the measurements are randomly chosen.
- **“Conjecture”**: Stable recovery from random measurements can be done using $C \dim(\mathcal{M} - \mathcal{M})$ samples if $\mathcal{M}$ is “nice”, say it is irreducible, or $C \dim(\mathcal{M} - \mathcal{M}) \log d$ samples if $\mathcal{M}$ is “not nice” (say many irreducible components).

Thank You!