List-Decodable Linear Regression

Sushrut Karmalkar* Adam Klivans* Pravesh Kothari†

*UT Austin
†CMU

This talk

Given: n samples as follows:

αn **inlier points** (x_i, y_i) s.t. $y_i = \langle x_i, \ell^* \rangle$; $\|\ell^*\|_2^2 = 1$ and $x_i \sim \mathcal{N}(0, I_{d\times d})$.

$(1-\alpha)n$ **outlier points** chosen arbitrarily and potentially adversarially.

Goal: Return a $O(1/\alpha)$ sized list L containing $\ell : \|\ell - \ell^*\| \leq 0.001$.

This talk

Given: n samples as follows:

αn **inlier points** (x_i, y_i) s.t. $\boxed{y_i = \langle x_i, \ell^* \rangle}$; $\|\ell^*\|_2^2 = 1$ and $x_i \sim \boxed{\mathcal{N}(0, I_{d\times d})}$.

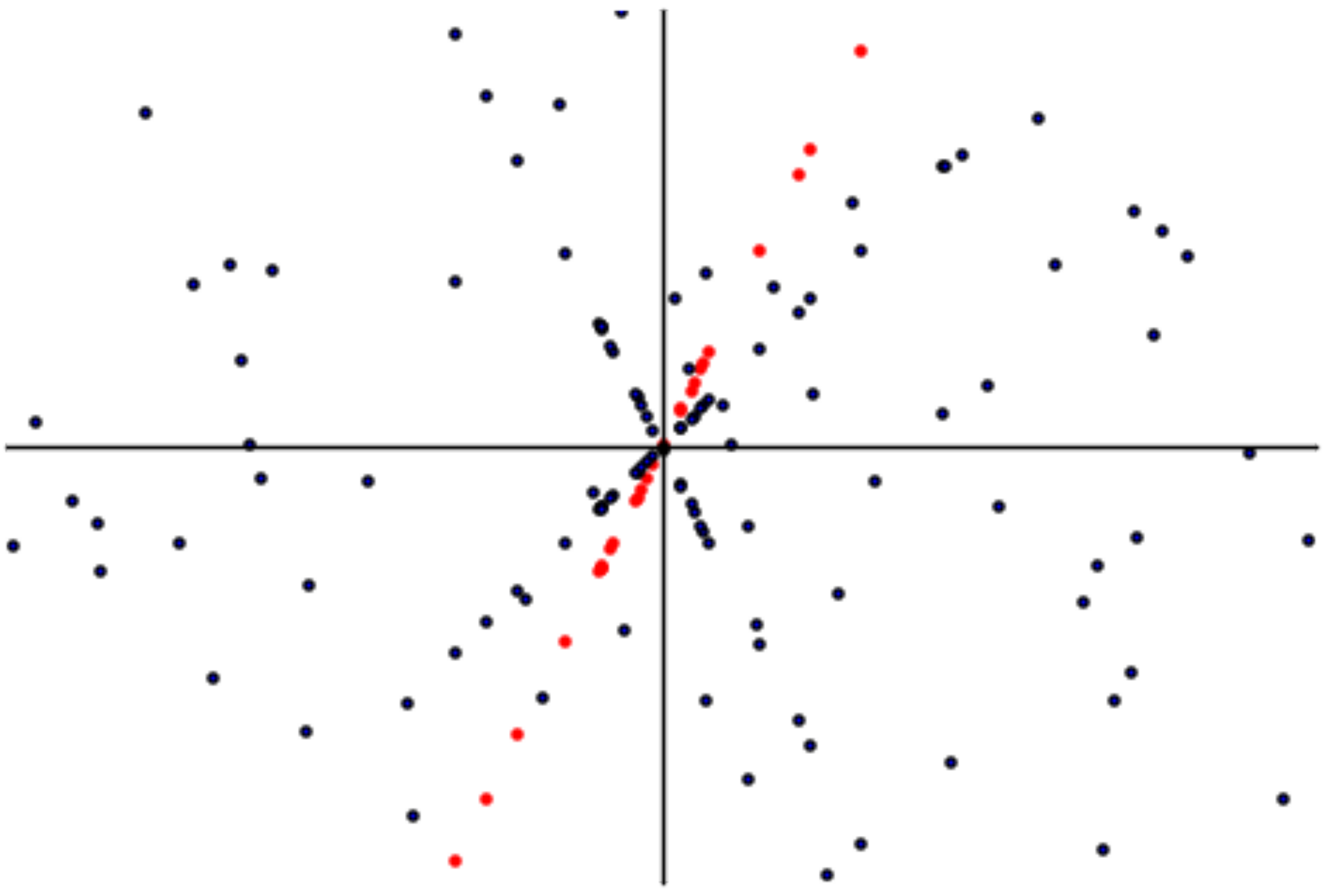
$(1-\alpha)n$ **outlier points** chosen arbitrarily and potentially adversarially.

Goal: Return a $O(1/\alpha)$ sized list L containing $\ell : \|\ell - \ell^*\| \leq 0.001$.

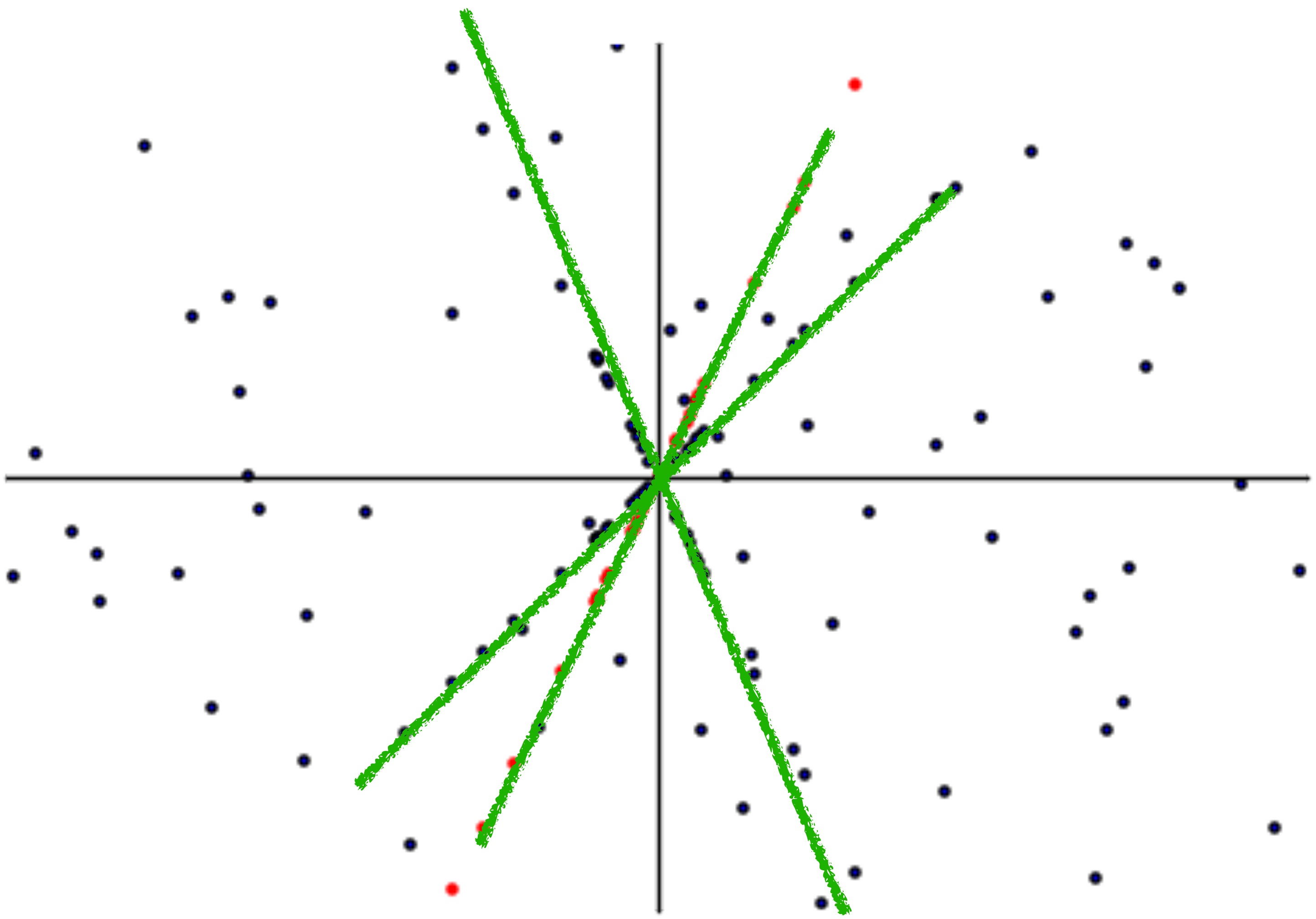
Certifiably anti-concentrated

Additive noise

This talk



This talk



Robust statistics

Classical Results: Statistical estimators that can recover signals after a fraction of the data is corrupted. [...'70s], ...

Robust statistics

Classical Results: Statistical estimators that can recover signals after a fraction of the data is corrupted. [...'70s], ...

2016 [<0.25 **adversarial** corruptions]**:** Computationally efficient estimators. [Lai-Rao-Vempala'16], [Diakonikolas-Kane-Kamath-Li-Moitra-Stewart'16], ...

Robust statistics

Classical Results: Statistical estimators that can recover signals after a fraction of the data is corrupted. […'70s], …

2016 [<0.25 **adversarial** corruptions]**:** Computationally efficient estimators. [Lai-Rao-Vempala'16], [Diakonikolas-Kane-Kamath-Li-Moitra-Stewart'16], …

2017 [>0.5 **adversarial** corruptions]**: List-decodable** mean estimation. [Charikar-Steinhart-Valiant'17]

Why List-decodable Regression?

Previous techniques: <0.25 adversarial corruptions [Klivans-Kothari-Meka'18], [Diakonikolas-Kamath-Kane-Li-Steinhardt-Stewart'18], [Prasad-Suggala-Balakrishnan-Ravikumar'18], [Diakonikolas-Kong-Stewart 19] …

Why List-decodable Regression?

Previous techniques: <0.25 adversarial corruptions [Klivans-Kothari-Meka'18], [Diakonikolas-Kamath-Kane-Li-Steinhardt-Stewart'18], [Prasad-Suggala-Balakrishnan-Ravikumar'18], [Diakonikolas-Kong-Stewart 19] …

List decodable regression → **Unique recovery for <0.5 adversarial corruptions**

Why List-decodable Regression?

Previous techniques: <0.25 adversarial corruptions [Klivans-Kothari-Meka'18], [Diakonikolas-Kamath-Kane-Li-Steinhardt-Stewart'18], [Prasad-Suggala-Balakrishnan-Ravikumar'18], [Diakonikolas-Kong-Stewart 19] …

List decodable regression → **Unique recovery for <0.5 adversarial corruptions**

Generalizes *Mixed* Linear regression: [Deveaux'89], [Jordan-Jacobs'94], … , [Li-Liang'18], …

Main Result

We have an algorithm that

Takes input: n samples with αn inliers $(1-\alpha)n$ outliers.

Returns: a list L of size $O(1/\alpha)$ s.t. $\exists\ \ell \in L$ s.t $\|\ell - \ell^*\|_2 \leq 0.001$ with probability at least 0.99

Time/Sample complexity: $d^{O(1/\alpha^8)}$.

Our list size is optimal.

[**K**-Klivans-Kothari'19], [Raghavendra-Yau'19]

Information Theoretic Lower Bounds

We show that list-decodable regression with a constant list size (or even $d-1$) is **impossible** when the **underlying distribution is not anti-concentrated**.

Surprisingly, this holds even for **uniform on hypercube**.

Open Problems

1. Which distributions other than the Gaussian are **certifiably anti-concentrated?**
2. Robust covariance estimation of a mixture of gaussians
   - [Bakshi-Kothari'19] Show an algorithm for robust subspace recovery using techniques very similar to ones used in this paper.

Thank You!