How to do theoretical research

Ankur Moitra (MIT)

FOCS October 28th, 2024

DISCLAIMER

Advice about research is like advice about running:

How do you run a marathon?

DISCLAIMER

Advice about research is like advice about running:

How do you run a marathon?

obvious	run hard, run long

DISCLAIMER

Advice about research is like advice about running:

How do you run a marathon?

obvious	run hard, run long
pragmatic	tie your shoes, hydrate well

DISCLAIMER

Advice about research is like advice about running:

How do you run a marathon?

obvious	run hard, run long
pragmatic	tie your shoes, hydrate well
preparation*	here’s a training regiment to follow

*this is what I’ll focus on

DISCLAIMER

There are many **styles** to research

DISCLAIMER

There are many **styles** to research

Useful to see other people’s blueprint, to help you craft your own

ASIDE

When I first started teaching, I co-taught with Tom Leighton

ASIDE

When I first started teaching, I co-taught with Tom Leighton

Incredible lecturer, **most of what I know about speaking I learned from him**

After lecture, he handed me his **script**

<u>Welcome</u> to 6.042/18.062 Discrete Math for CS (more affectionately known as Proofs, Proofs & more Proofs!). It probably won't take you long to figure out how the course got this nickname. ☺

3

30

⋮

After lecture, he handed me his **script**

Welcome to 6.042/18.062 Discrete Math for CS (more affectionately known as Proofs, Proofs & more Proofs!). It probably won't take you long to figure out how the course got this nickname. ☺

3

30

My name is Tom Leighton and I'll be co-teaching this term with **Ankur Moitra and Debayan Gupta.** **

⋮

After lecture, he handed me his **script**

Welcome to 6.042/18.062 Discrete Math for CS (more affectionately known as Proofs, Proofs & more Proofs!). It probably won't take you long to figure out how the course got this nickname. ☺

3

30

My name is Tom Leighton and I'll be co-teaching this term with **Ankur Moitra and Debayan Gupta.** **

Stand up.

After lecture, he handed me his **script**

Welcome to 6.042/18.062 Discrete Math for CS (more affectionately known as Proofs, Proofs & more Proofs!). It probably won't take you long to figure out how the course got this nickname. ☺

3

30

My name is Tom Leighton and I'll be co-teaching this term with **Ankur Moitra and Debayan Gupta.** **

Stand up.

For a while, I just thought that's how you teach

After lecture, he handed me his **script**

Welcome to 6.042/18.062 Discrete Math for CS (more affectionately known as Proofs, Proofs & more Proofs!). It probably won’t take you long to figure out how the course got this nickname. ☺

3

30

My name is Tom Leighton and I’ll be co-teaching this term with **Ankur Moitra and Debayan Gupta.** **

Stand up.

For a while, I just thought that’s how you teach

But it’s not my style --- same can be true about research

ADVICE

My most important advice:

Work on good problems

ADVICE

My most important advice:

Work on good problems

What makes a good problem? And how do you find them?

COMMON STRATEGIES

Strategy #1: What your advisor tells you to work on

COMMON STRATEGIES

Strategy #1: What your advisor tells you to work on

+ They know the community, what people care about

COMMON STRATEGIES

Strategy #1: What your advisor tells you to work on

- + They know the community, what people care about
- + They know the literature, can better predict what is true and what is within reach

COMMON STRATEGIES

Strategy #1: What your advisor tells you to work on

+ They know the community, what people care about

+ They know the literature, can better predict what is true and what is within reach

− You do the best research when you **own** the problem

COMMON STRATEGIES

Strategy #2: Open question from recent papers in your area

COMMON STRATEGIES

Strategy #2: Open question from recent papers in your area

+ Usually concrete and well-defined

COMMON STRATEGIES

Strategy #2: Open question from recent papers in your area

- + Usually concrete and well-defined
- + Already know where to look for relevant techniques

COMMON STRATEGIES

Strategy #2: Open question from recent papers in your area

+ Usually concrete and well-defined

+ Already know where to look for relevant techniques

– People state open questions for all sorts of reasons, not always the right ones

COMMON STRATEGIES

Strategy #2: Open question from recent papers in your area

+ Usually concrete and well-defined

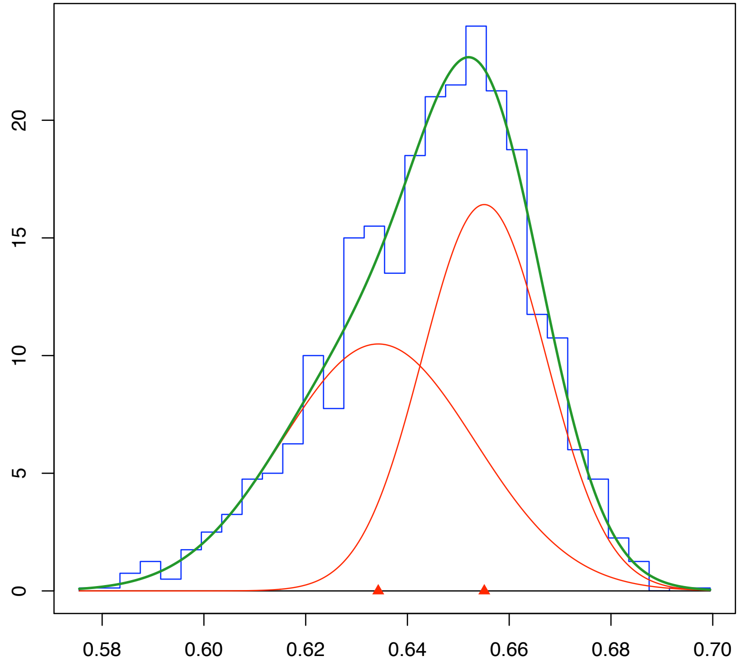
+ Already know where to look for relevant techniques

− People state open questions for all sorts of reasons, not always the right ones

Some problems are more difficult than they are interesting, need to learn how to tell

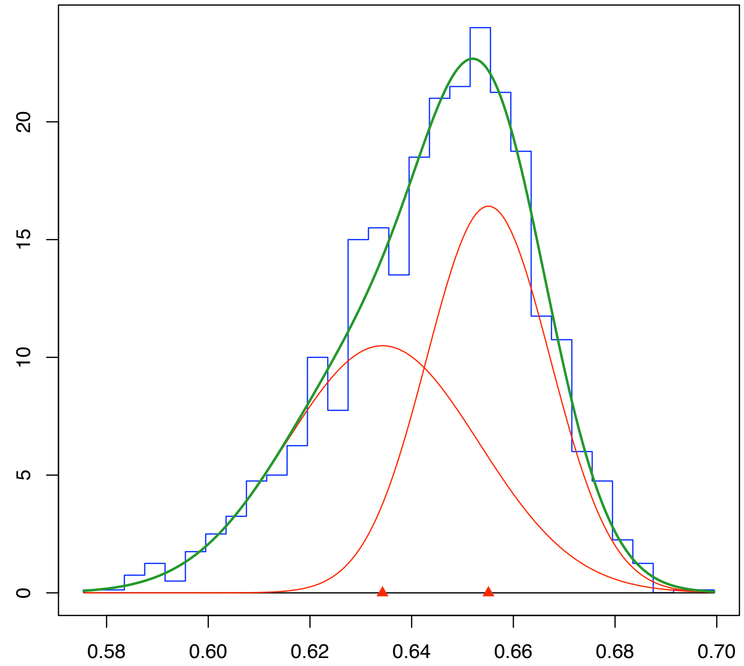
RESEARCH STORIES

In Summer 2009, Adam Kalai, Greg Valiant and I were working on learning mixtures of Gaussians



RESEARCH STORIES

In Summer 2009, Adam Kalai, Greg Valiant and I were working on learning mixtures of Gaussians



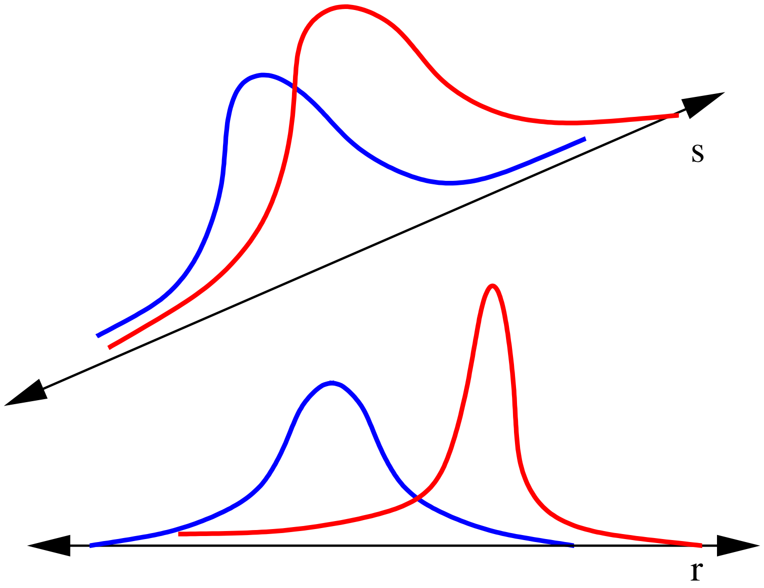
Can we learn the components from samples?

TECHNICAL ASIDE

This was an old question, but we had a new angle…

TECHNICAL ASIDE

This was an old question, but we had a new angle



Strategy: Learn the high-dimensional parameters through one-dimensional projections

TECHNICAL ASIDE

But what was known about the one-dimensional problem? Surely this had to have been solved already

TECHNICAL ASIDE

But what was known about the one-dimensional problem? Surely this had to have been solved already

Statistical question: Why can't two mixtures with different parameters yield almost identical distributions?

TECHNICAL ASIDE

But what was known about the one-dimensional problem? Surely this had to have been solved already

For any two mixtures:

IF their parameters are ε-different

THEN they must be $\delta(\varepsilon)$-far as distributions

If $\delta(\varepsilon)$ is a polynomial in ε and the dimension d, we say the model is **polynomially identifiable**

TECHNICAL ASIDE

Any algorithm must prove this along the way

TECHNICAL ASIDE

Any algorithm must prove this along the way

Notions like asymptotic efficiency sidestep it, by taking the limit as the number of samples goes to infinity

BACK TO THE STORY

In Fall 2010, Dan Spielman invited me to come to Yale to talk about it in their Statistics Seminar

James Dwight Dana House

BACK TO THE STORY

In Fall 2010, Dan Spielman invited me to come to Yale to talk about it in their Statistics Seminar

James Dwight Dana House

Turned into an hours long debate about why our results weren't already known

TAKEAWAY

The definitions in a field color the way it sees the world

TAKEAWAY

The definitions in a field color the way it sees the world

What new perspectives do you bring?

TAKEAWAY

The definitions in a field color the way it sees the world

What new perspectives do you bring?

For me, it's about algorithms, **not just about what they can do but how the problem ought to be formulated**

ADVICE #1

Be deliberate about what perspectives you bring

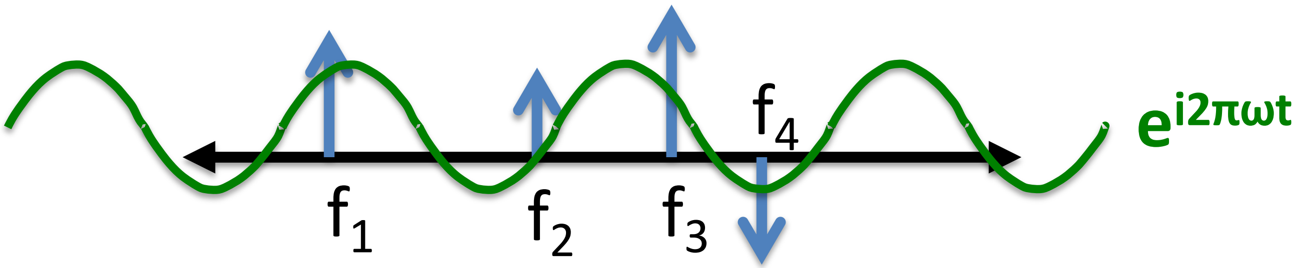
ADVICE #1

Be deliberate about what perspectives you bring

Your research journey is less random than it might seem

EPILOGUE

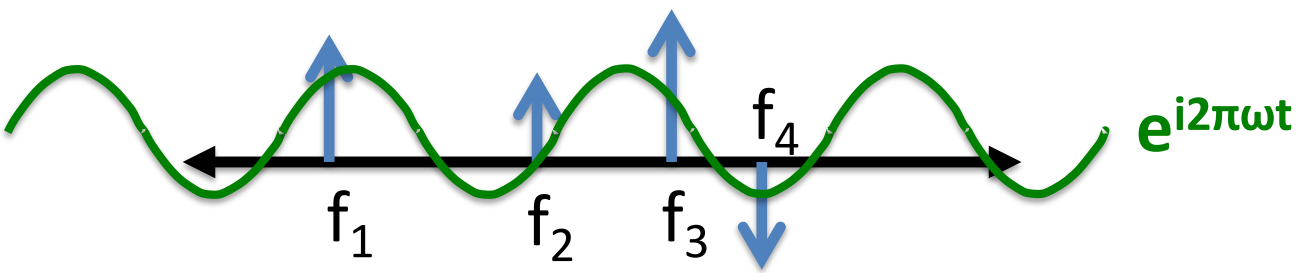
In Spring 2014, I learned about super-resolution from Carlos Fernandez-Granda



Goal: Recover the locations of particles, with only low frequency measurements

EPILOGUE

In Spring 2014, I learned about super-resolution from Carlos Fernandez-Granda

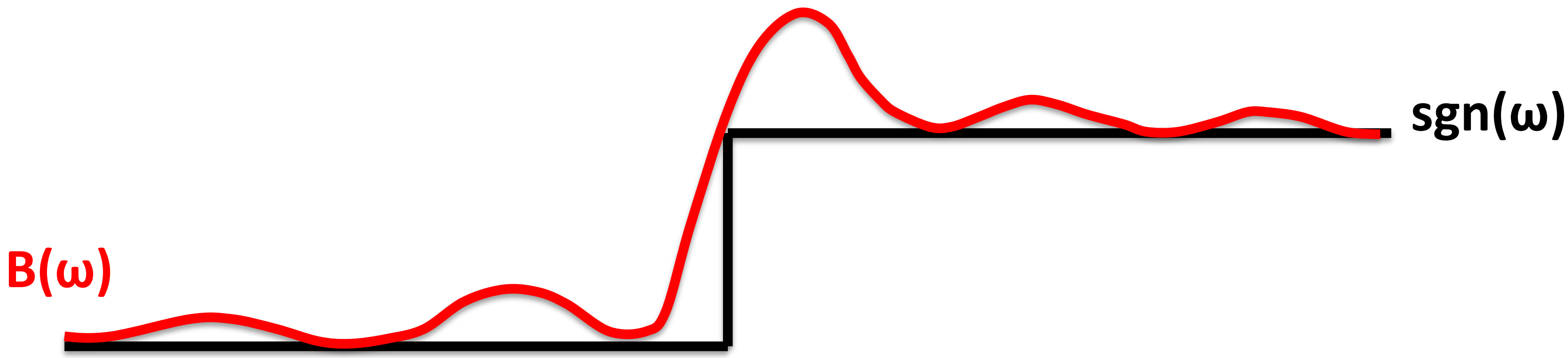


Goal: Recover the locations of particles, with only low frequency measurements

Existing work assumed no noise, but are the locations robustly identifiable?

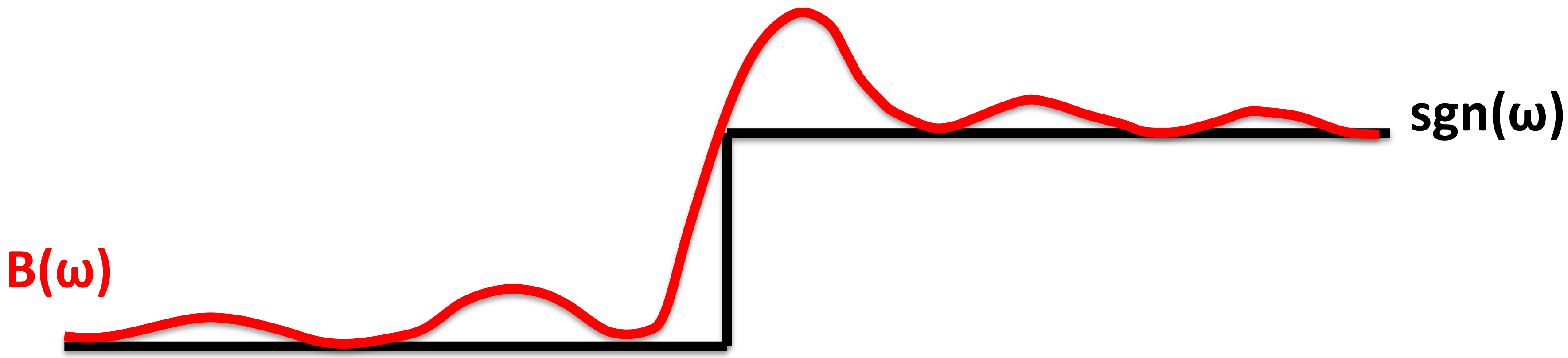
EPILOGUE

It turns out this is related to condition numbers of Vandermonde matrices and **extremal functions**



EPILOGUE

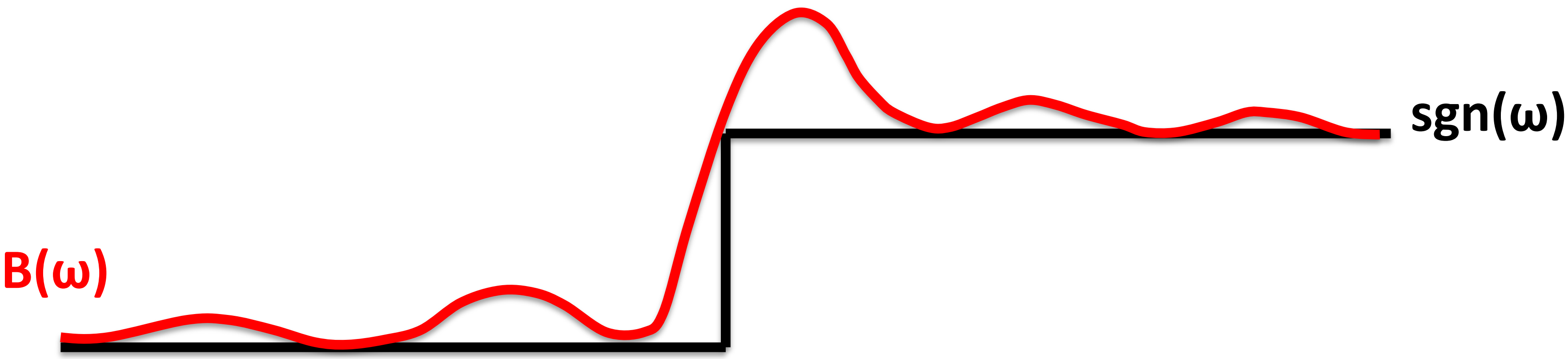
It turns out this is related to condition numbers of Vandermonde matrices and **extremal functions**



“Luck is what happens when preparation meets opportunity”

EPILOGUE

It turns out this is related to condition numbers of Vandermonde matrices and **extremal functions**



"~~Luck~~ is what happens when preparation meets opportunity"

Research

ADVICE #2

Find the easiest problem you can't solve

ADVICE #2

Find the easiest problem you can't solve

Bothersome question: Our algorithm for learning mixtures of Gaussians is brittle. Can it be made robust?

RESEARCH STORIES

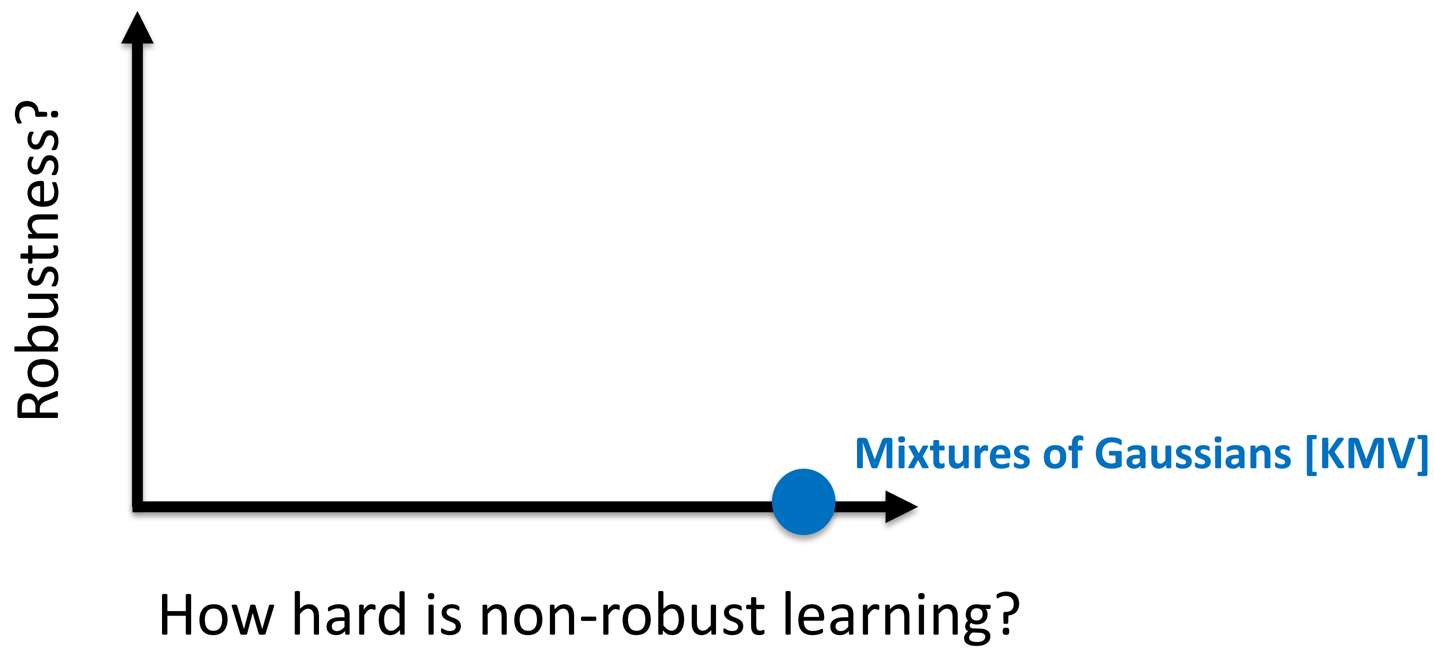
In Fall 2015, Jerry Li and Gautam Kamath asked me:

Can you robustly learn a single Gaussian?

RESEARCH STORIES

In Fall 2015, Jerry Li and Gautam Kamath asked me:

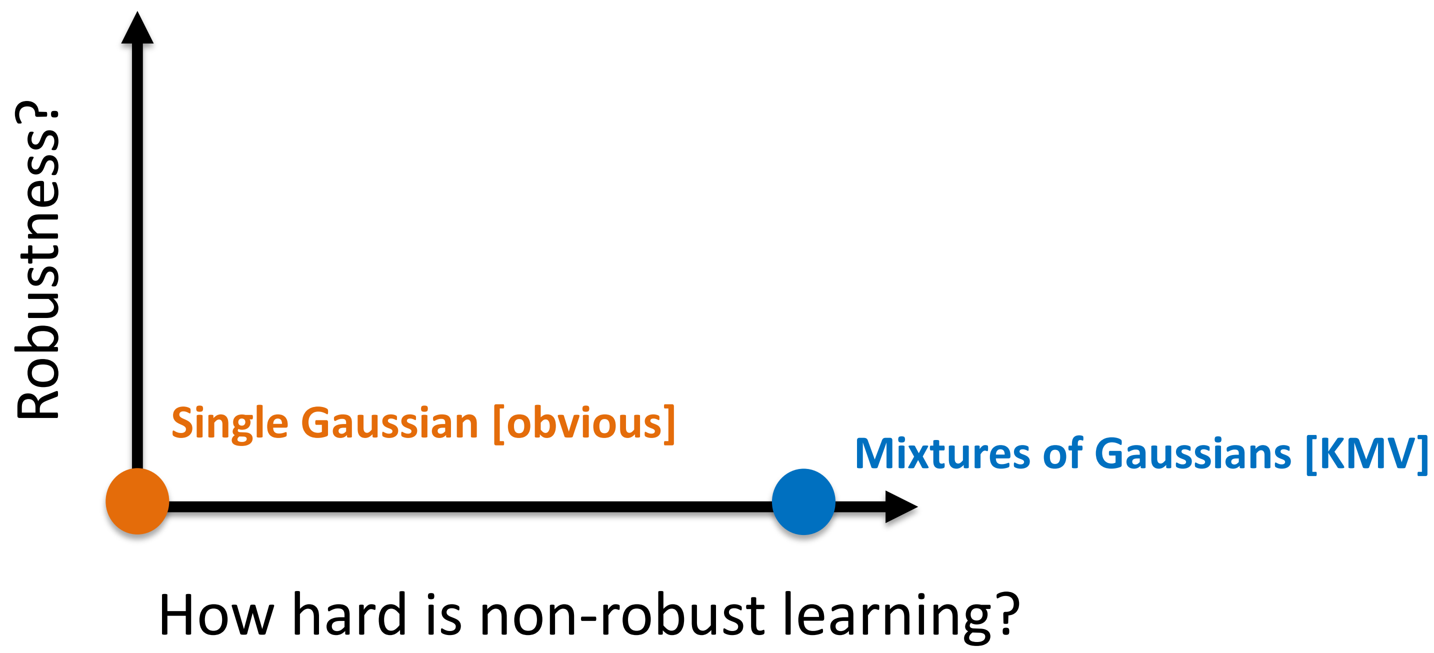
Can you robustly learn a single Gaussian?



RESEARCH STORIES

In Fall 2015, Jerry Li and Gautam Kamath asked me:

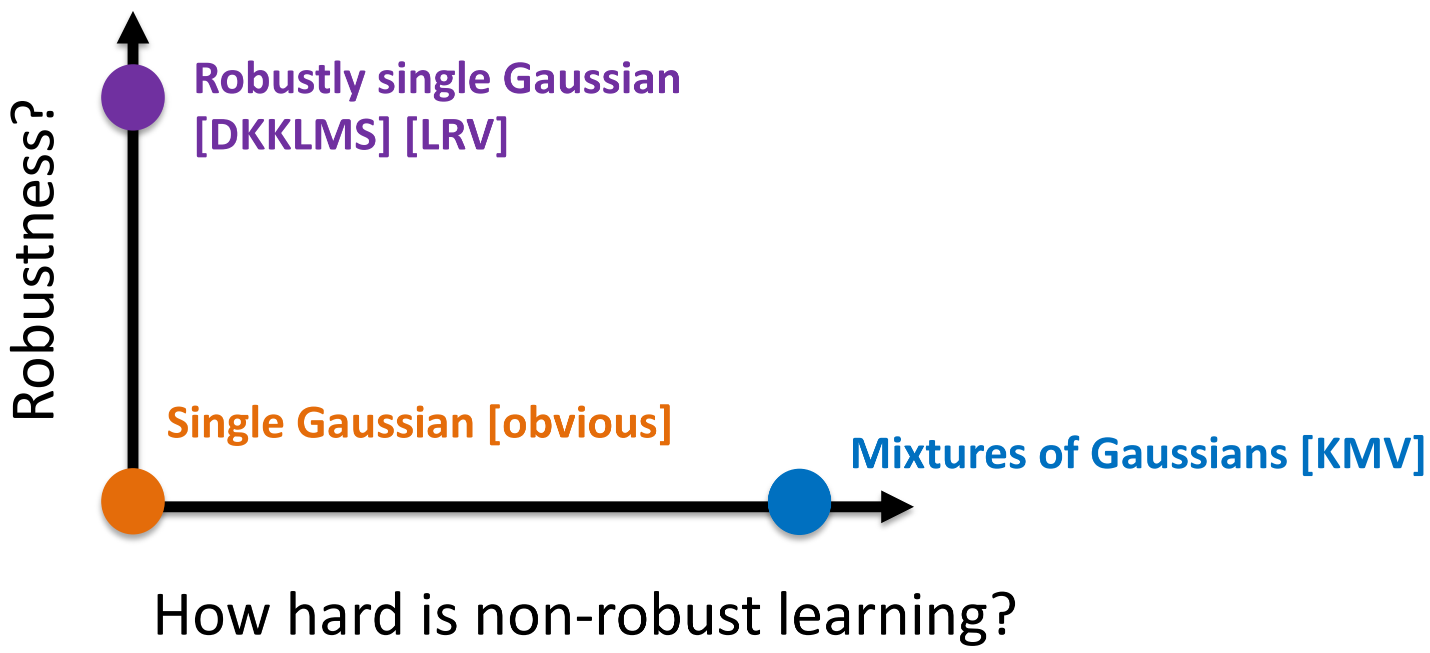
Can you robustly learn a single Gaussian?



RESEARCH STORIES

In Fall 2015, Jerry Li and Gautam Kamath asked me:

Can you robustly learn a single Gaussian?



EPILOGUE

Subsequently many works **handling more errors via list decoding**, **giving lower bounds against statistical query algorithms**, **weakening the distributional assumptions**, **exploiting sparsity**, **working with more complex generative models**

EPILOGUE

Subsequently many works **handling more errors via list decoding**, **giving lower bounds against statistical query algorithms**, **weakening the distributional assumptions**, **exploiting sparsity**, **working with more complex generative models**

Eventually we got robust learning for mixtures of Gaussians --- needed full arsenal of tools, weren't ready for this earlier

ADVICE #2

Find the easiest problem you can't solve

ADVICE #2

Find the easiest problem you can’t solve

The more embarrassing, the better!

ADVICE #3

Don’t be afraid to reason by analogy

RESEARCH STORIES

I learned about linear dynamical systems from Moritz Hardt and Tengyu Ma

arXiv > cs > arXiv:1609.05191

Computer Science > Machine Learning

[Submitted on 16 Sep 2016 (v1), last revised 11 Feb 2019 (this version, v2)]

Gradient Descent Learns Linear Dynamical Systems

Moritz Hardt, Tengyu Ma, Benjamin Recht

RESEARCH STORIES

I learned about linear dynamical systems from Moritz Hardt and Tengyu Ma

arXiv > cs > arXiv:1609.05191

Computer Science > Machine Learning

[Submitted on 16 Sep 2016 (v1), last revised 11 Feb 2019 (this version, v2)]

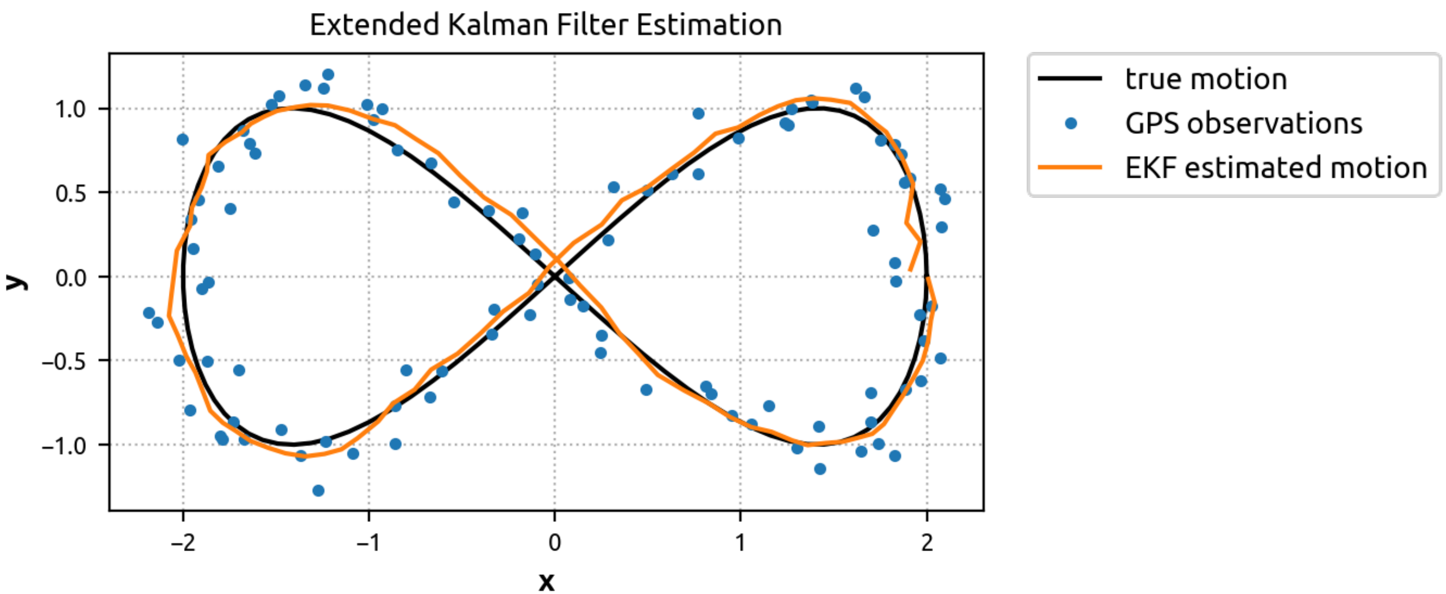
Gradient Descent Learns Linear Dynamical Systems

Moritz Hardt, Tengyu Ma, Benjamin Recht

Sounds nice, but I didn't have anything to contribute

RESEARCH STORIES

In Fall 2021, I finally understood what a linear dynamical system is



The observations are a Gaussian process, with particular algebraic structure

RESEARCH STORIES

If the method of moments works for mixtures of Gaussians, can it work for linear dynamical systems too?

RESEARCH STORIES

If the method of moments works for mixtures of Gaussians, can it work for linear dynamical systems too?

Ainesh Bakshi, Allen Liu, Morris Yau and I gave a simple and powerful algorithm which removes many unneeded assumptions

ADVICE #3

Don’t be afraid to reason by analogy

ADVICE #3

Don’t be afraid to reason by analogy

When you come with your own map, much easier to explore off the beaten path

ADVICE #4

Know the landscape inside out

ADVICE #4

Know the landscape inside out

Are there neighboring problems that are hard? How can you exploit the differences?

ADVICE #4

Know the landscape inside out

Are there neighboring problems that are hard? How can you exploit the differences?

Are there neighboring problems that are easy? Can those techniques be a starting point?

ADVICE #5

Learn lots of powerful tools

ADVICE #5

Learn lots of powerful tools

Go to talks, read papers, but first ask yourself: Are you surprised?

ADVICE #5

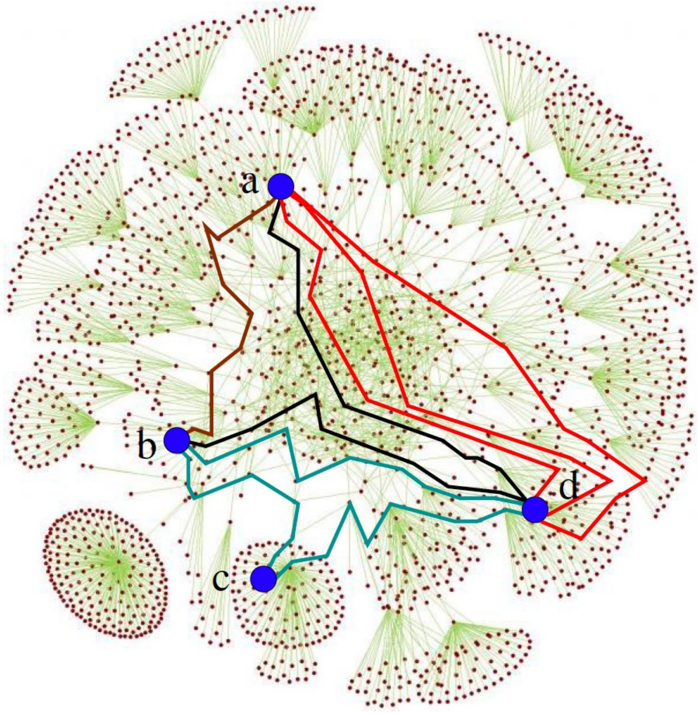
Learn lots of powerful tools

Go to talks, read papers, but first ask yourself: Are you surprised?

If you are, then get to the bottom of it

RESEARCH STORIES

My PhD thesis introduced vertex sparsifiers



Can approximate all flows and cuts between terminals on a much smaller graph

RESEARCH STORIES

When I told my advisor, he didn't believe me

"Oh yeah, what's the vertex sparsifier for this?"

RESEARCH STORIES

When I told my advisor, he didn't believe me

"Oh yeah, what's the vertex sparsifier for this?"

"I don't know"

RESEARCH STORIES

When I told my advisor, he didn't believe me

"Oh yeah, what's the vertex sparsifier for this?"

"I don't know"

"Let's make it easier, what about for this?"

RESEARCH STORIES

When I told my advisor, he didn't believe me

"Oh yeah, what's the vertex sparsifier for this?"

"I don't know"

"Let's make it easier, what about for this?"

"I don't know"

RESEARCH STORIES

When I told my advisor, he didn't believe me

> **"But I don't have to construct it, just show you can't prove me wrong!"**

RESEARCH STORIES

When I told my advisor, he didn't believe me

"But I don't have to construct it, just show you can't prove me wrong!"

My proof was through zero-sum games

RESEARCH STORIES

When I told my advisor, he didn't believe me

> **"But I don't have to construct it, just show you can't prove me wrong!"**

My proof was through zero-sum games

I think understanding duality is a superpower, allows you to view challenging problems in different ways

ADVICE #5

Learn lots of powerful tools

ADVICE #5

Learn lots of powerful tools

Powerful tools don’t always need to be fancy, but you need to know them deeply

ADVICE #6

Find the right level of simplification

ADVICE #6

Find the right level of simplification

These vary by subfield, e.g. expander decompositions

"Let's just start with expanders and see where that gets us…"

ADVICE #6

Find the right level of simplification

These vary by subfield, e.g. expander decompositions

"Let's just start with expanders and see where that gets us…"

But keep a research journal so you know where you are in the expedition

ADVICE #7

Work with great people

ADVICE #7

Work with great people

Greatly expands the scope of problems you're equipped to work on

ADVICE #7

Work with great people

Greatly expands the scope of problems you're equipped to work on

Noga Alon, Matthew Andrews, Sanjeev Arora, Ainesh Bakshi, Afonso Bandeira, Boaz Barak, Aditya Bhaskara, Mark Braverman, Guy Bresler, Victor-Emmanuel Brunel, Moses Charikar, Sitan Chen, Zongchen Chen, Diego Cifuentes, Costis Daskalakis, Michelle Delcourt, Ilias Diakonikolas, Anindya De, Cole Franks, Rong Ge, Ran Gelles, Noah Golowich, Yoni Halpern, MohammadTaghi Hajiaghayi, Yoni Halpern, Linus Hamilton, Moritz Hardt, Sam Hopkins, Nicole Immorlica, Saachi Jain, Adam Kalai, Gautam Kamath, Daniel Kane, Ravi Kannan, Howard Karloff, Jon Kelner, Younhun Kim, Frederic Koehler, Pravesh Kothari, Hannah Lawrence, Holden Lee, Tom Leighton, Jerry Li, Shi Li, Allen Liu, Kuikui Liu, Brendan Lucier, Tengyu Ma, Aleks Madry, Nitya Mani, Raghu Meka, David Mimno, Elchanan Mossel, Ryan O'Donnell, Chirag Pabbaraju, Guillem Perarnau, Amelia Perry, Luke Postle, Andrew Postlewaite, Aaron Potechin, Oded Regev, Andrej Risteski, Dhruv Rohatgi, Prasad Raghavendra, Oded Regev, Philippe Rigollet, Sushant Sachdeva, Amit Sahai, Mike Saks, Colin Sandon, Rocco Servedio, Anish Sevekari, David Sontag, David Steurer, Benny Sudakov, Ewin Tang, Moshe Tennenholtz, Luca Trevisan, John Urschel, Greg Valiant, Aravindan Vijayaraghavan, Alex Wein, John Wright, Yichen Wu, Morris Yau, Michael Zhu

Summary:

- Many strategies and styles for finding good problems
  Don’t be afraid to look for your own!
- Learn the landscape of your area --- reasoning by analogy can be helpful
- Hone your instincts for finding the right level of abstraction

Thanks!

Summary:

- Many strategies and styles for finding good problems
  Don’t be afraid to look for your own!
- Learn the landscape of your area --- reasoning by analogy can be helpful
- Hone your instincts for finding the right level of abstraction

Happy to take questions