Nonnegative Polynomials: from Optimization to Control and Learning

Amir Ali Ahmadi
Princeton, ORFE
Affiliated member of PACM, COS, MAE, CSML

IEOR-DRO Seminar, Columbia University
March 2019

PRINCETON UNIVERSITY ORFE

Optimization over nonnegative polynomials

Defn. A polynomial $p(x) := p(x_1, \ldots, x_n)$ is nonnegative if $p(x) \geq 0, \forall x \in \mathbb{R}^n$.

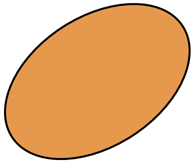
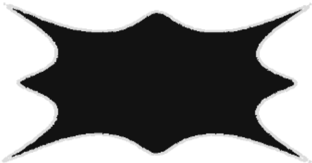
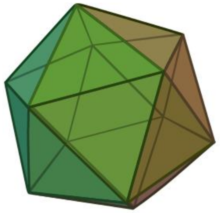
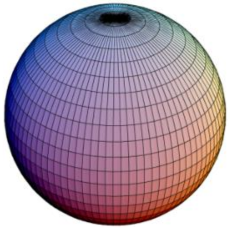
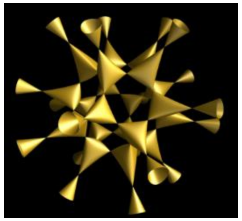
Example: When is

$$p(x_1, x_2) = c_1 x_1^4 - 6x_1^3 x_2 - 4x_1^3 + c_2 x_1^2 x_2^2 + 10x_1^2 + 12x_1 x_2^2 + c_3 x_2^4$$

nonnegative?
nonnegative over a given basic semialgebraic set?

Basic semialgebraic set: $\{x \in \mathbb{R}^n | \; g_i(x) \geq 0\}$

Ex:
$$x_1^3 - 2x_1 x_2^4 \geq 0$$
$$x_1^4 + 3x_1 x_2 - x_2^6 \geq 0$$

Why do this?!

Applications in optimization and statistics

Is $p(x) \geq 0$ on $\{g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$?

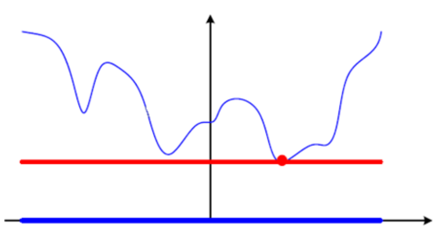
Polynomial optimization

- (Tight) lower bounds on polynomial optimization problems

$$\min_x p(x)$$
$$\text{s.t.} \quad g_i(x) \geq 0$$

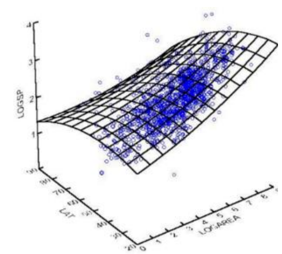
$\updownarrow$

$$\max_\gamma \gamma$$
$$\text{s.t.} \quad p(x) - \gamma \geq 0,$$
$$\forall x \in \{g_i(x) \geq 0\}$$



Shaped-constrained regression

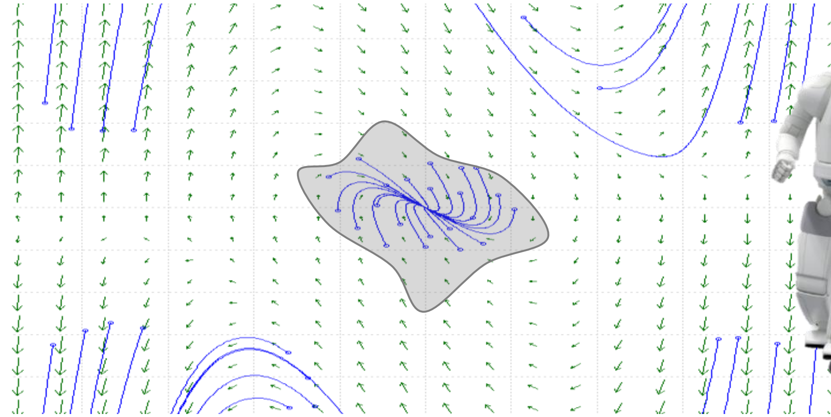
- Fitting a polynomial to data subject to shape constraints (e.g., convexity, or monotonicity)



$$\frac{\partial p(x)}{\partial x_j} \geq 0, \forall x \in B$$

Applications in analysis of dynamical systems

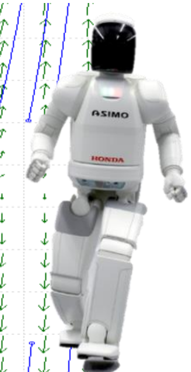
$$\dot{x} = f(x) \qquad f : \mathbb{R}^n \to \mathbb{R}^n$$

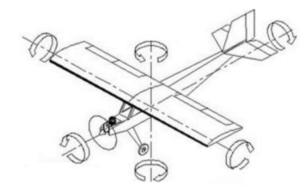


Ex. $\dot{x}_1 = -x_2 + \frac{3}{2}x_1^2 - \frac{1}{2}x_1^3 x_2$

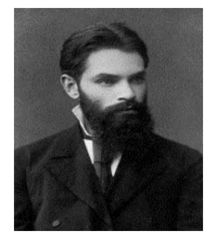
$\dot{x}_2 = 3x_1 - x_1 x_2$

Locally asymptotic stability (LAS) of equilibrium points



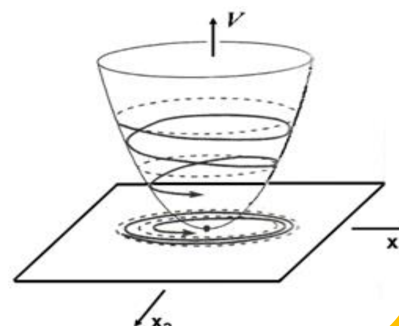


Lyapunov's theorem (and its converse):

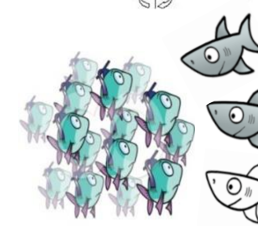


The origin is LAS if and only if there exists a C^1 function $V: \mathbb{R}^n \to \mathbb{R}$ that vanishes at the origin and a scalar $\beta > 0$ such that

$$V(x) > 0$$
$$V(x) \leq \beta \Rightarrow \dot{V}(x) = \nabla V(x)^T f(x) < 0$$



(Also implies that $\{x \mid V(x) \leq \beta\}$ is in the region of attraction (ROA))



- Stability of equilibrium prices in economics
- Convergence analysis of algorithms, …

Outline of the rest of the talk

1) Algebraic certificates of nonnegativity

An optimization-free Positivstellensatz (w/ G. Hall)

2) Local asymptotic stability of large-scale systems

DSOS and SDSOS Optimization (w/ A. Majumdar)

3) Learning dynamical systems with side information

A new application of nonnegative polynomials (w/ B. El Khadir)

Part 1: Algebraic certificates of nonnegativity

$$\begin{aligned} p(x) &= x_1^4 - 6x_1^3x_2 + 2x_1^3x_3 + 6x_1^2x_3^2 + 9x_1^2x_2^2 - 6x_1^2x_2x_3 \\ &\quad -14x_1x_2x_3^2 + 4x_1x_3^3 + 5x_3^4 - 7x_2^2x_3^2 + 16x_2^4 \end{aligned}$$

Nonnegative

$\Uparrow$

SOS

$$\begin{aligned} p(x) &= (x_1^2 - 3x_1x_2 + x_1x_3 + 2x_3^2)^2 + (x_1x_3 - x_2x_3)^2 \\ &\quad +(4x_2^2 - x_3^2)^2. \end{aligned}$$

Important fact:

- Optimization over sum of squares (SOS) polynomials is a semidefinite program (SDP)!

A polynomial p of degree $2d$ is SOS if and only if $\exists Q \succcurlyeq 0$ such that

$$p(x) = z(x)^T Q z(x)$$

where $z = [1, x_1, \ldots, x_n, x_1x_2, \ldots, x_n^d]^T$ is the vector of monomials of degree up to d.

Positivstellensätze

$p(x) > 0, \forall x \in \mathbb{R}^n$

Artin

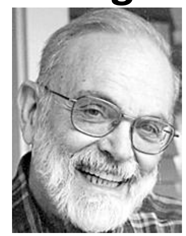
If $p(x) \geq 0, \forall x \in \mathbb{R}^n$,
then $\exists$ SOS q s.t. $p \cdot q$ SOS

1927

20th century

1974

Stengle



1991

1993

$p(x) > 0$,
$\forall x \in S = \{x \mid g_i(x) \geq 0\}$

Schmudgen

If $p(x) > 0, \forall x \in S$,
then $p(x) = \sigma_0(x) + \sum_i \sigma_i(x) g_i(x) + \sum_{ij} \sigma_{ij}(x) g_i(x) g_j(x) + \cdots$, with $\sigma_0, \sigma_i, \ldots$ SOS

(Requires some compactness assumptions)

Putinar

If $p(x) > 0, \forall x \in S$,
then $p(x) = \sigma_0(x) + \sum_i \sigma_i(x) g_i(x)$,
where σ_0, σ_i are SOS

Search for these SOS polynomials (when degree is fixed) requires SDP.

An optimization-free Positivstellensatz (1/2)

$$p(x) > 0, \forall x \in \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \dots, m\}$$

$2d$ =maximum degree of p, g_i

$\Updownarrow$ Under compactness assumptions, i.e., $\{x \mid g_i(x) \geq 0\} \subseteq B(0, R)$

$\exists\, r \in \mathbb{N}$ such that

$$\left(f(v^2 - w^2) - \frac{1}{r}\left(\sum_i (v_i^2 - w_i^2)^2\right)^d + \frac{1}{2r}\left(\sum_i (v_i^4 + w_i^4)\right)^d\right) \cdot \left(\sum_i v_i^2 + \sum_i w_i^2\right)^{r^2}$$

has **nonnegative coefficients**,

where f is a form in $n + m + 3$ variables and of degree $4d$, which can be explicitly written from p, g_i and R.

[AAA, Hall, *Math of OR*] (2018 INFORMS Young Researchers Prize)

PRINCETON UNIVERSITY ORFE

An optimization-free Positivstellensatz (2/2)

$$p(x) > 0 \text{ on } \{x \mid g_i(x) \geq 0\} \Leftrightarrow$$
$$\exists r \in \mathbb{N} \text{ s.t. } \left(f(v^2 - w^2) - \frac{1}{r}\left(\sum_i (v_i^2 - w_i^2)^2\right)^d + \frac{1}{2r}\left(\sum_i (v_i^4 + w_i^4)\right)^d \right) \cdot \left(\sum_i v_i^2 + \sum_i w_i^2\right)^{r^2}$$
has ≥ 0 coefficients

- $p(x) > 0$ on $\{x \mid g_i(x) \geq 0\} \Leftrightarrow f$ is pd
- **Result by Polya (1928):**

f **even** and pd $\Rightarrow \exists r \in \mathbb{N}$ such that $f(z) \cdot \left(\sum_i z_i^2\right)^r$ has nonnegative coefficients.

- Make $f(z)$ even by considering $f(v^2 - w^2)$. We lose positive definiteness of f with this transformation.
- Add the positive definite term $\frac{1}{2r}\left(\sum_i (v_i^4 + w_i^4)\right)^d$ to $f(v^2 - w^2)$ to make it positive definite. **Apply Polya's result.**
- The term $-\frac{1}{r}\left(\sum_i (v_i^2 - w_i^2)^2\right)^d$ ensures that the converse holds as well.

Corollary:

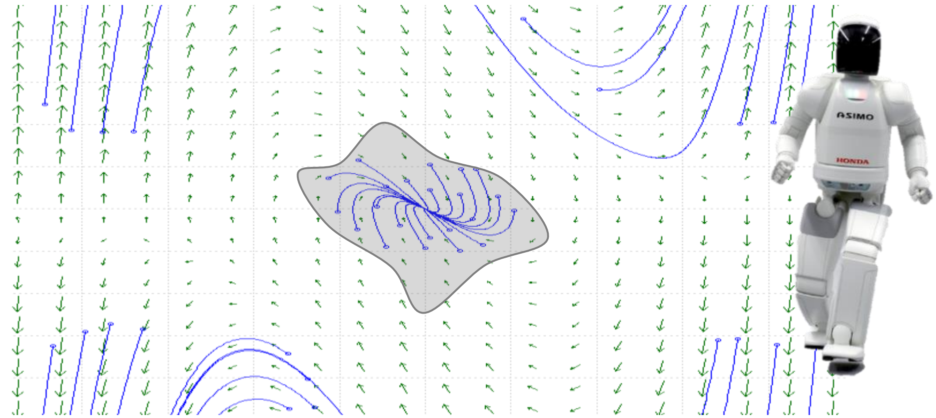
LP/SOCP-based converging hierarchies for polynomial optimization.
(e.g. based on "DSOS or SDSOS polynomials")

Part 2: Verifying local asymptotic stability (LAS)

$$\dot{x} = f(x) \qquad f : \mathbb{R}^n \to \mathbb{R}^n$$

Ex. $\dot{x}_1 = -x_2 + \frac{3}{2}x_1^2 - \frac{1}{2}x_1^3 x_2$

$\dot{x}_2 = 3x_1 - x_1 x_2$



Lyapunov's theorem (and its converse):

The origin is LAS if and only if there exists a C^1 function $V: \mathbb{R}^n \to \mathbb{R}$ that vanishes at the origin and a scalar $\beta > 0$ such that

$$V(x) > 0$$
$$V(x) \leq \beta \Rightarrow \dot{V}(x) = \nabla V(x)^T f(x) < 0$$

(Also implies that $\{x | \, V(x) \leq \beta\}$ is in the (ROA))

$\deg(f) = 1$: Linear systems

$$\dot{x} = Ax$$

- LAS decidable in polynomial time.
- Iff eigenvalues of A in the open left half complex plane.
- Iff there exists a **quadratic Lyapunov function**!

What if deg(*f*)>1? …

From linear to nonlinear

What if deg(f)>1? …

- **Conjecture of Arnol'd (1976): LAS is undecidable** (still open)

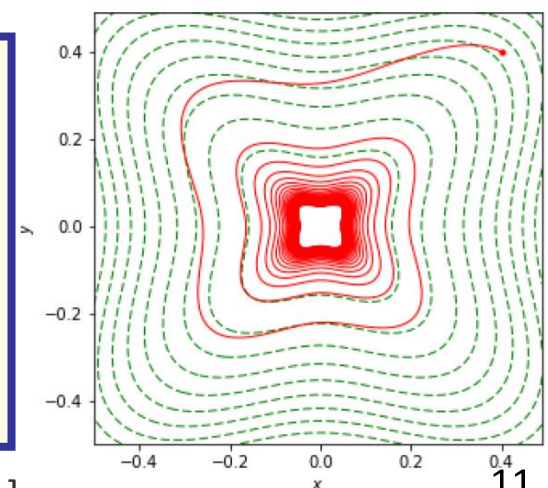
Existence of a **polynomial Lyapunov function**, together with a **computable upper bound** on its degree would imply decidability (e.g., by quantifier elimination).

Thm: Deciding local asymptotic stability of cubic vector fields is **strongly NP-hard**. This is true even for gradient vector fields.

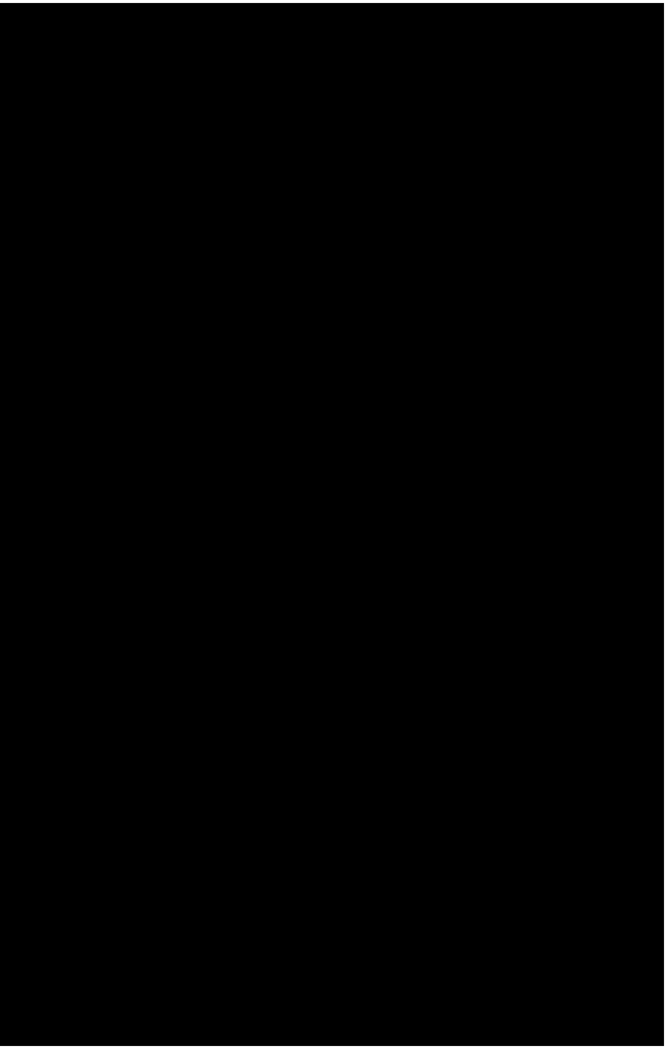
[AAA, *American Control Conference*]

Thm: The origin of the following vector field is LAS but the **there is no polynomial Lyapunov function** (of any degree):

$$f(x,y) = \begin{pmatrix} -2y(-x^4+2x^2y^2+y^4) \\ 2x(x^4+2x^2y^2-y^4) \end{pmatrix} - (x^2+y^2)\begin{pmatrix} 2x(x^4+2x^2y^2-y^4) \\ 2y(-x^4+2x^2y^2+y^4) \end{pmatrix}$$



Practical challenges ≠ theoretical challenges!



Controller designed by SOS

Since the year 2000 (Parrilo's PhD thesis), SOS optimization has been widely used for verifying LAS (via polynomial Lyapunov functions) and for designing stabilizing polynomial control laws.

Pros:

- Deals with nonlinear systems directly.
- Gives formal proofs of stability in a fully-automated fashion.

Cons:

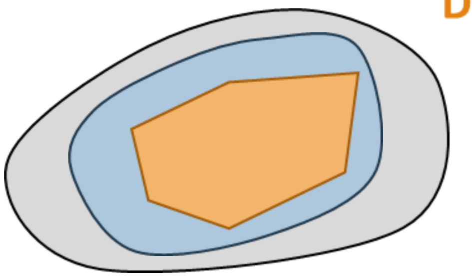
- Scalability! Often limited to systems with few state variables.
- Leads to *expensive semidefinite programming constraints*.

[Majumdar, AAA, Tedrake] (Best paper award - *IEEE Conf. on Robotics and Automation*)

dsos and sdsos polynomials (1/2)

Sum of squares **(sos)**	$p(x) = z(x)^T Qz(x), Q \succcurlyeq 0$	**SDP**

PSD cone$:= \{Q \mid Q \succcurlyeq 0\}$



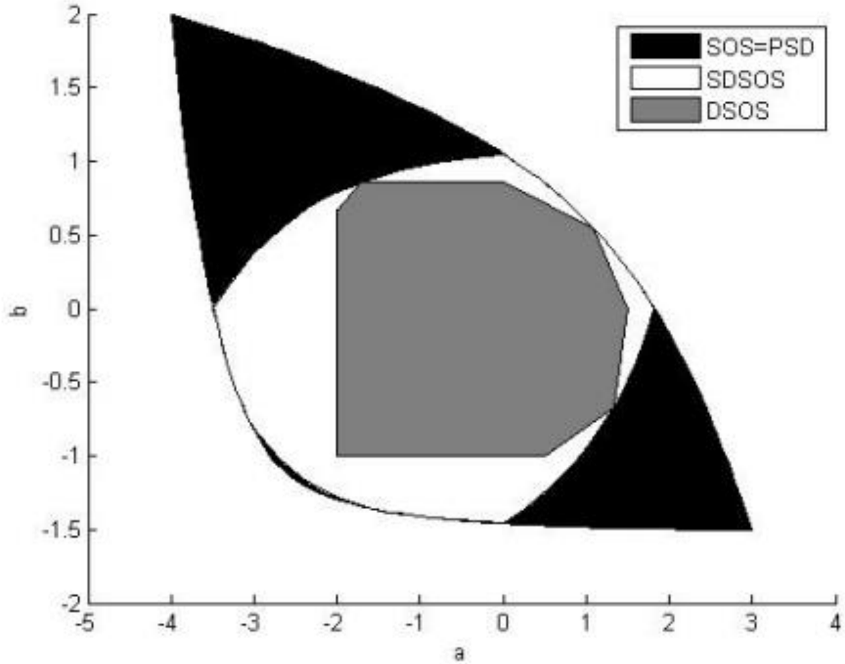
DD cone $:= \{Q \mid Q_{ii} \geq \sum_{j \neq i} |Q_{ij}|, \forall i\}$

SDD cone $:= \{Q \mid \exists$ **diagonal** D **with** $D_{ii} > 0$ **s.t.** $DQD\ dd\}$

Diagonally dominant sum of squares **(dsos)**	$p(x) = z(x)^T Qz(x), Q\ diagonally\ dominant\ (\mathrm{dd})$	**LP**

Scaled diagonally dominant sum of squares **(sdsos)**	$p(x) = z(x)^T Qz(x), Q\ scaled\ diagonally\ dominant\ (\mathrm{sdd})$	**SOCP**

dsos and sdsos polynomials (2/2)



$$p(x_1, x_2) = x_1^4 + x_2^4 + a x_1^3 x_2 + (1 - \frac{1}{2}a - \frac{1}{2}b) x_1^2 x_2^2 + 2b x_1 x_2^3$$

Can do better by changing basis…

dsos/sdsos + change of basis (1/2)

$$p(x) = x_1^4 - 6x_1^3x_2 + 2x_1^3x_3 + 6x_1^2x_3^2 + 9x_1^2x_2^2 - 6x_1^2x_2x_3 - 14x_1x_2x_3^2 + 4x_1x_3^3 + 5x_3^4 - 7x_2^2x_3^2 + 16x_2^4$$

$$p(x) = z^T(x)Qz(x)$$

$$Q = \begin{pmatrix} 1 & -3 & 0 & 1 & 0 & 2 \\ -3 & 9 & 0 & -3 & 0 & -6 \\ 0 & 0 & 16 & 0 & 0 & -4 \\ 1 & -3 & 0 & 2 & -1 & 2 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 2 & -6 & 4 & 2 & 0 & 5 \end{pmatrix}$$

$$z(x) = (x_1^2, x_1x_2, x_2^2, x_1x_3, x_2x_3, x_3^2)^T$$

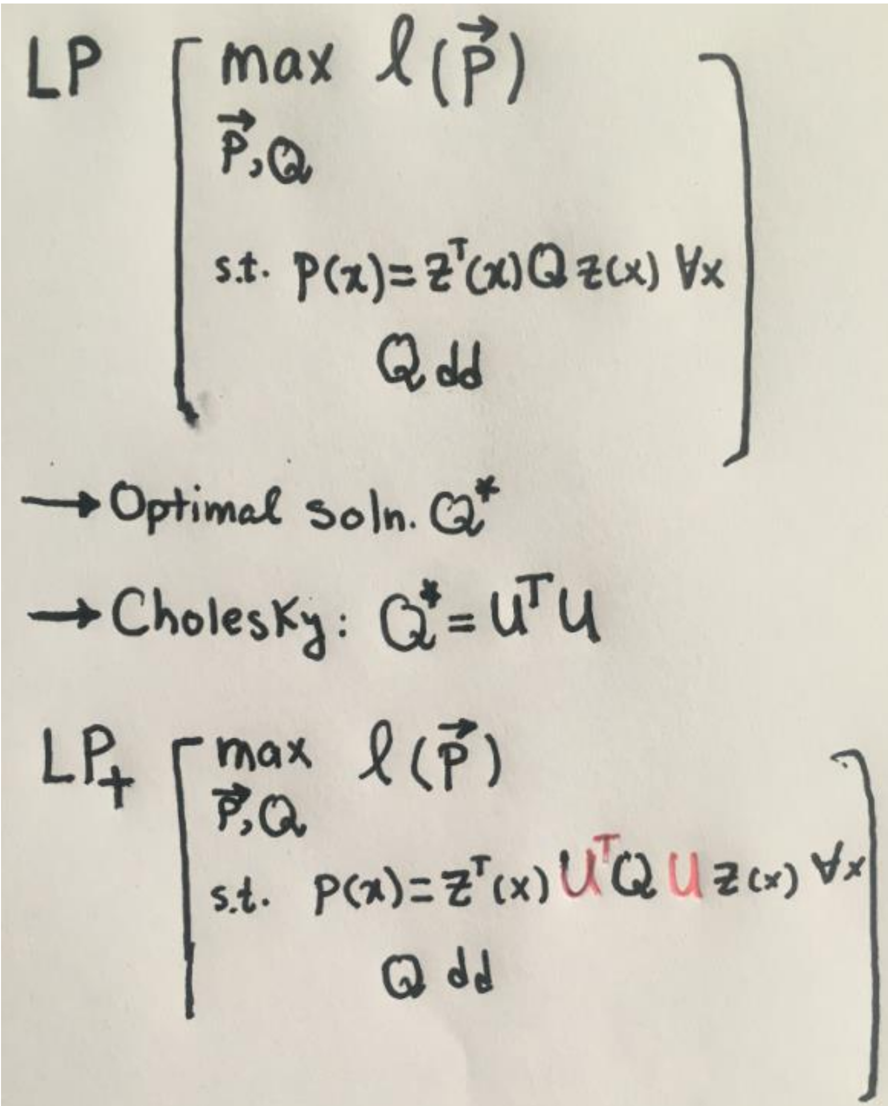
$$p(x) = \tilde{z}^T(x) \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4. \end{pmatrix} \tilde{z}(x)$$

$$\tilde{z}(x) = \begin{pmatrix} 2x_1^2 - 6x_1x_2 + 2x_1x_3 + 2x_3^2 \\ x_1x_3 - x_2x_3 \\ x_2^2 - \frac{1}{4}x_3^2 \end{pmatrix}$$

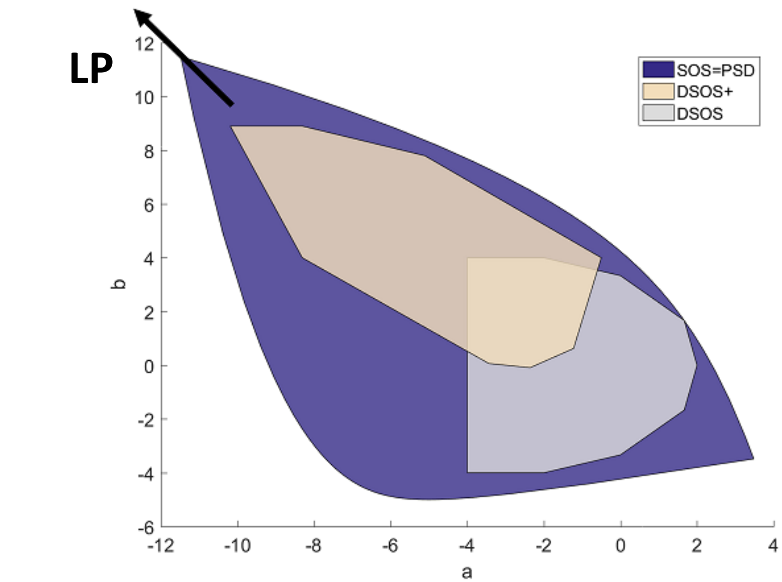
Goal: iteratively improve $z(x)$

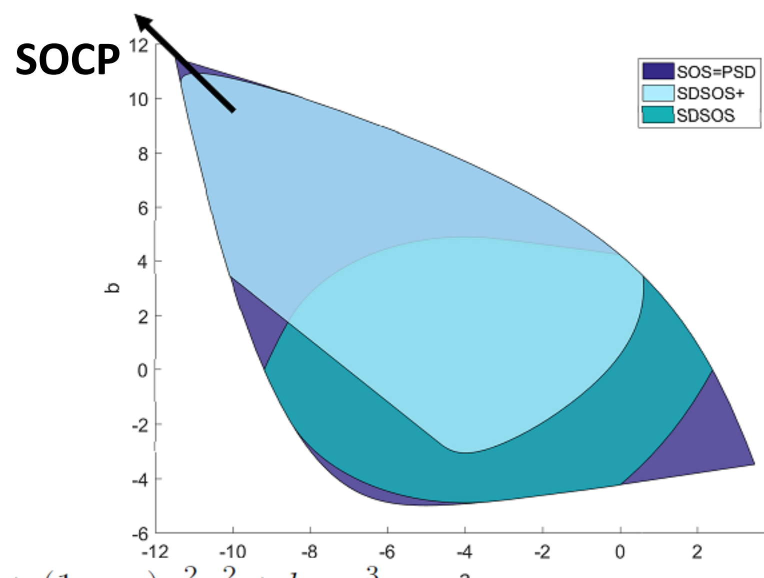
[AAA, Hall, *Contemporary Mathematics*]

dsos/sdsos + change of basis (2/2)



Works beautifully!





$$p_{a,b}(x_1, x_2) = 2x_1^4 + 2x_2^4 + a x_1^3 x_2 + (1-a) x_2^2 x_2^2 + b x_1 x_2^3$$

PRINCETON UNIVERSITY ORFE

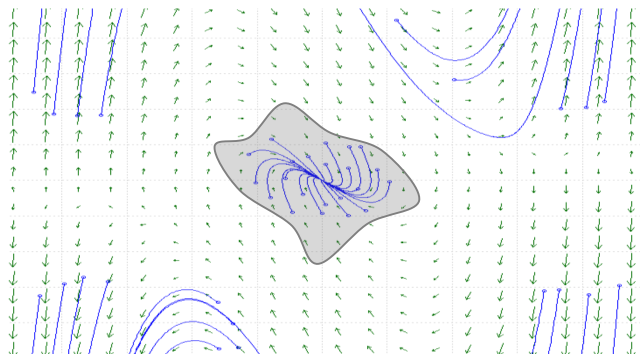
Applications in control

(see paper for applications in statistics, finance, combinatorial and polynomial optimization)

Reminder

$\dot{x} = f(x, u)$

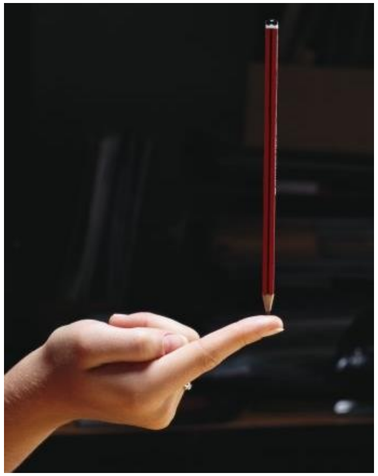
Stability of equilibrium points



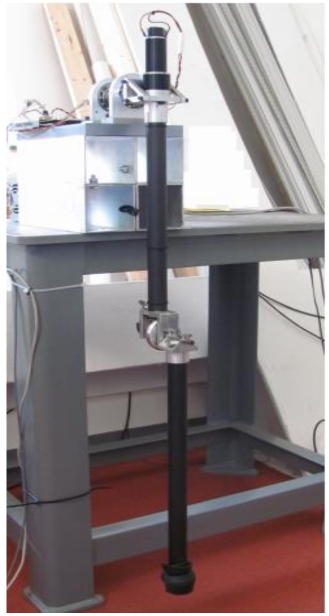
$$V(x) > 0,$$
$$V(x) \le \beta \Rightarrow \dot{V}(x) < 0$$

implies $\{x \mid V(x) \le \beta\}$ is in the region of attraction (ROA)

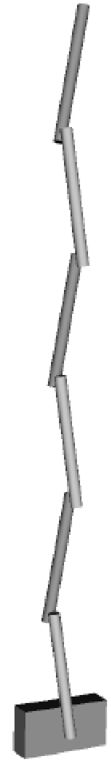
Stabilizing the inverted N-link pendulum (2N states)



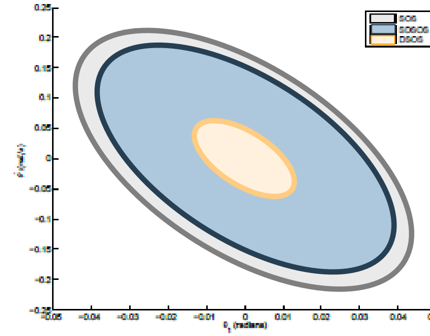
N=1



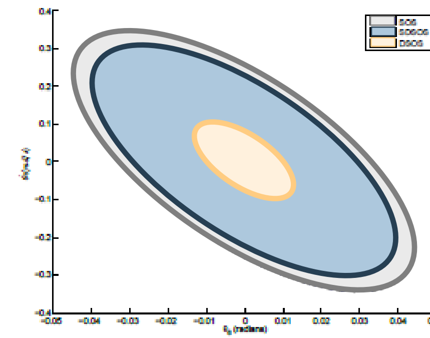
N=2



N=6



(a) θ_1-$\dot{\theta}_1$ subspace.



(b) θ_6-$\dot{\theta}_6$ subspace.

Runtime:

2N (# states)	4	6	8	10	12	14	16	18	20	22
DSOS	< 1	0.44	2.04	3.08	9.67	25.1	74.2	200.5	492.0	823.2
SDSOS	< 1	0.72	6.72	7.78	25.9	92.4	189.0	424.74	846.9	1275.6
SOS (SeDuMi)	< 1	3.97	156.9	1697.5	23676.5	∞	∞	∞	∞	∞
SOS (MOSEK)	< 1	0.84	16.2	149.1	1526.5	∞	∞	∞	∞	∞

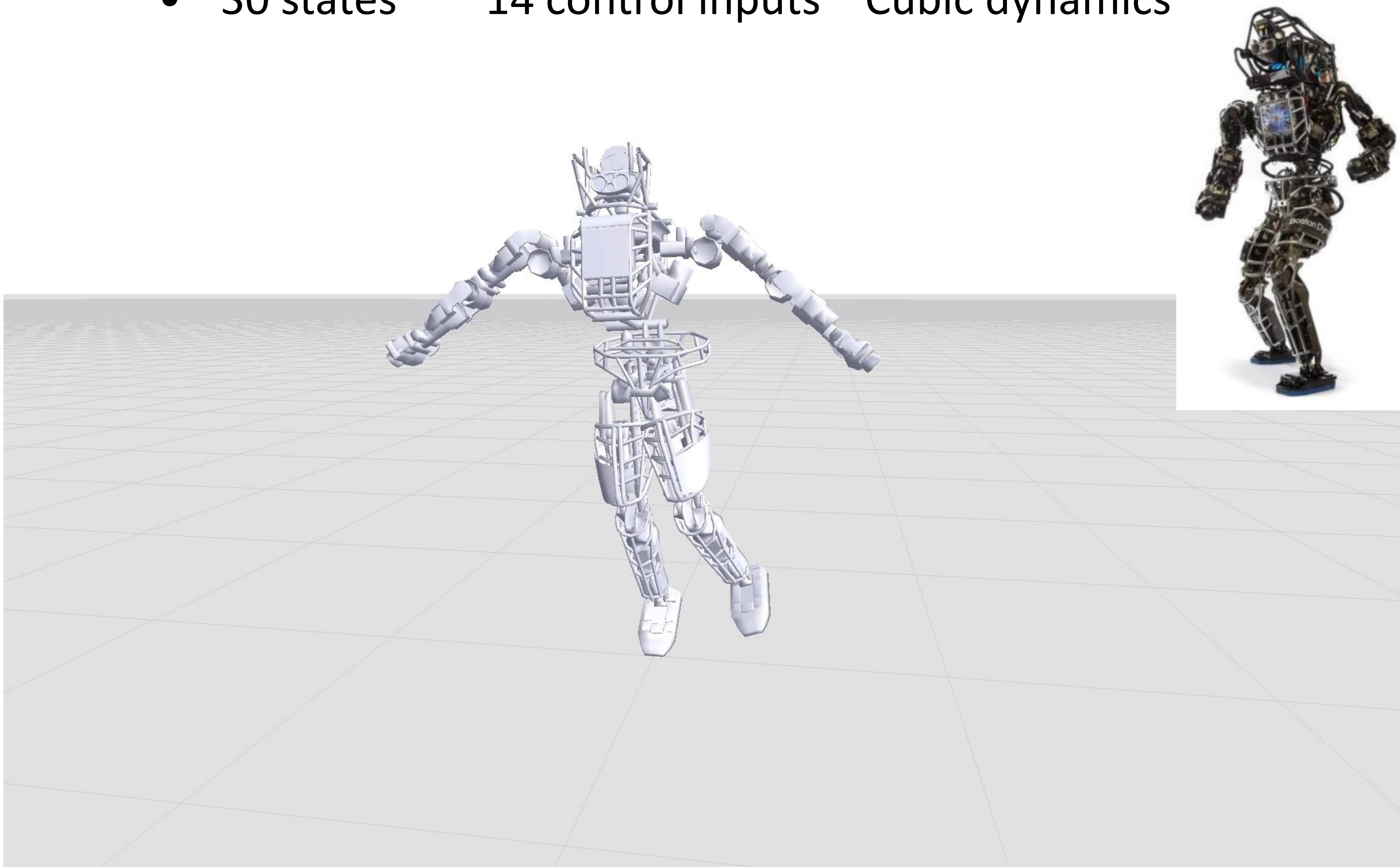
ROA volume ratio:

2N (states)	4	6	8	10	12
ρ_{dsos}/ρ_{sos}	0.38	0.45	0.13	0.12	0.09
ρ_{sdsos}/ρ_{sos}	0.88	0.84	0.81	0.79	0.79

[Majumdar, AAA, Tedrake, CDC]

Stabilizing ATLAS

- 30 states 14 control inputs Cubic dynamics



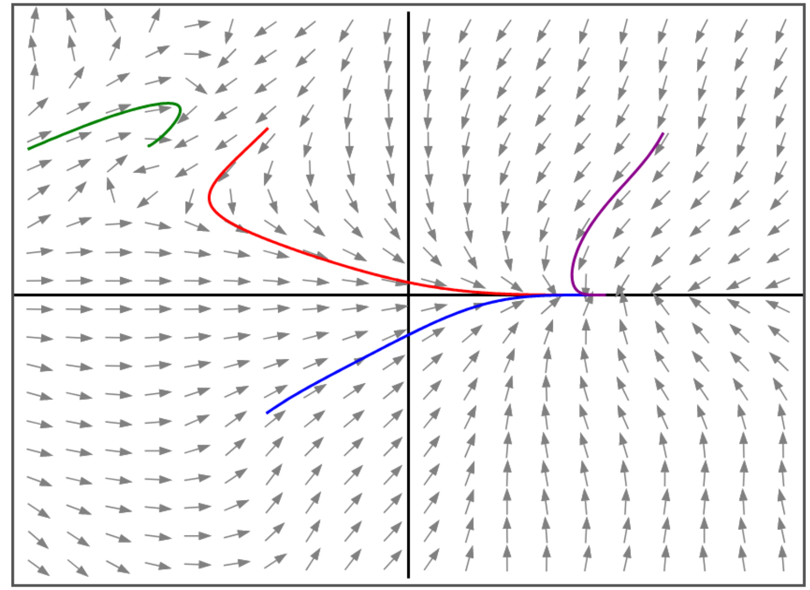
Done by **SDSOS Optimization**

[Majumdar, AAA, Tedrake, CDC]

PRINCETON UNIVERSITY ORFE

https://github.com/spot-toolbox/spotless

Part 3: Learning dynamical systems with side information



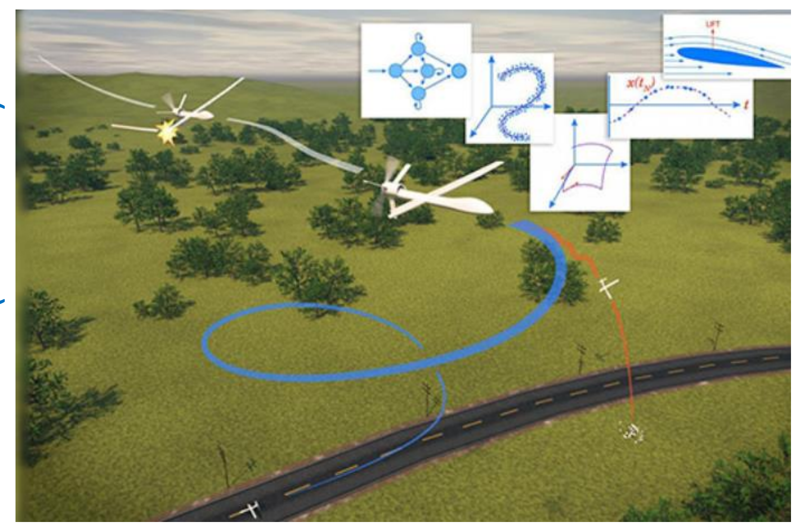
- Goal is to learn a dynamical system

  $$\dot{x} = f(x) \quad (\text{where } f: \mathbb{R}^n \to \mathbb{R}^n)$$

  from a *limited* number of *noisy* measurements of its trajectories.

- Data is scarce and expensive.
- Side information (physical laws/contextual information) needs to be exploited to enable learning.

MURI (2019-2024)



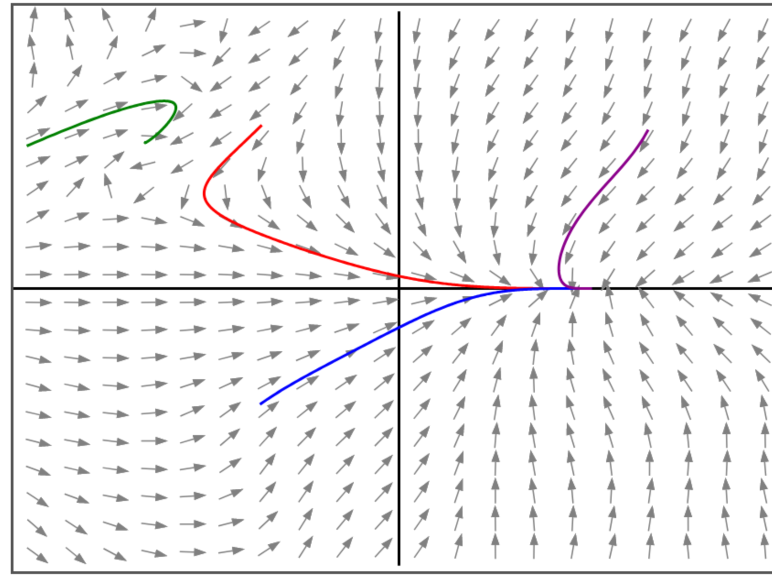
PRINCETON UNIVERSITY ORFE

A more precise formulation

- Goal is to learn a dynamical system

  $\dot{x} = f(x)$ (where $f: \mathbb{R}^n \to \mathbb{R}^n$)

  from a *limited* number of *noisy* measurements of its trajectories.



Examples of "side information":

- Equilibrium points (and their stability)
- Invariance of certain sets
- Decrease of certain energy functions
- Sign conditions on derivatives of states
- Monotonicity conditions
- Incremental stability
- (Non)reachability of a set B from a set A
- …

- Parametrize a polynomial vector field $p: \mathbb{R}^n \to \mathbb{R}^n$.
- Use SOS optimization to impose side information as constraints on p.
- Pick the p that best explains the data.

(Related prior work: Shaped-constrained regression [AAA, Curmei, Hall]; cf. Chap. 7 of PhD thesis of Hall)

A concrete epidemiology example

A model from the epidemiology literature for spread of Gonorrhea in a heterosexual population:

$$\dot{x} = f_1(x,y) = -a_1 x + b_1(1-x)y$$

$$\dot{y} = f_2(x,y) = -a_2 y + b_2(1-y)x$$

$x(t)$: fraction of infected males at time t

$y(t)$: fraction of infected females at time t

a_1: recovery rate of males

a_2: recovery rate of females

b_1: infection rate of males

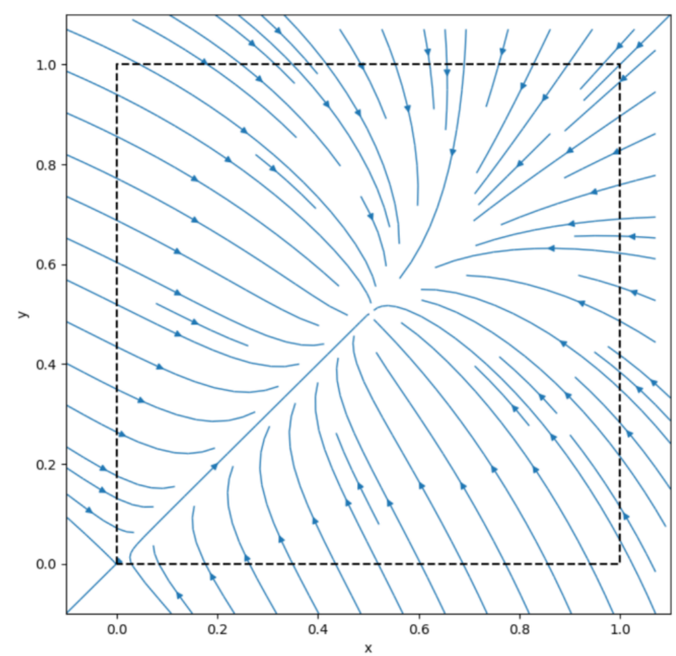
b_2: infection rate of females

For our experiments:

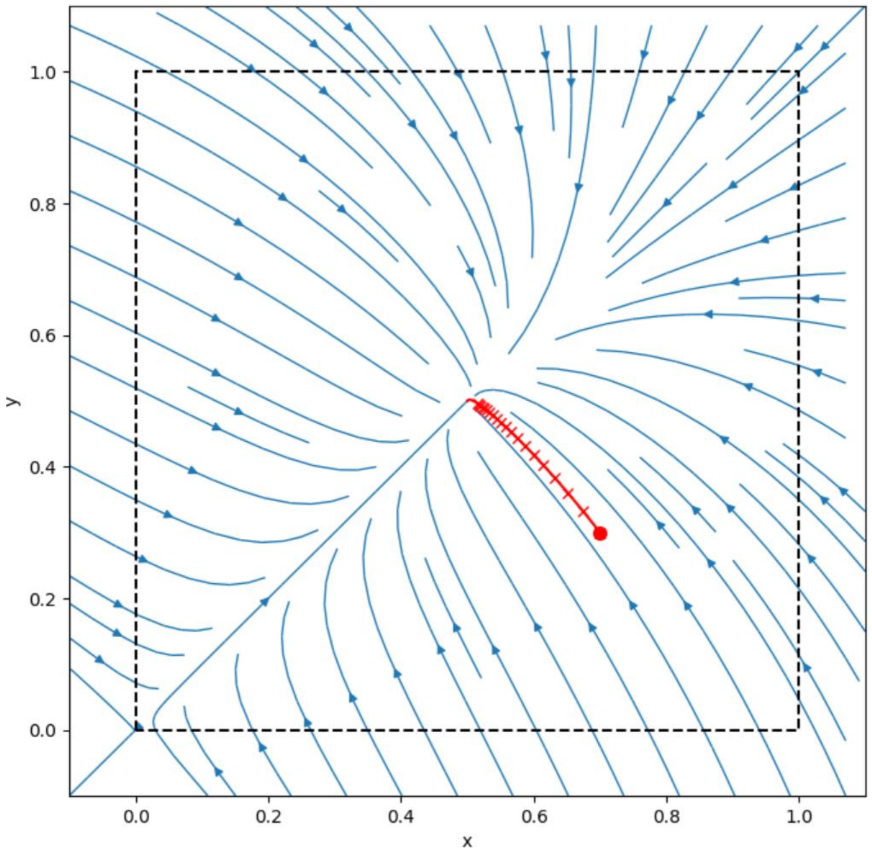
$a_1 = a_2 = .1; b_1 = b_2 = .05.$

This is taken to be "the ground truth".

- The dynamics (both its parameters and its special structure) is unknown to us.
- We only get to observe noisy trajectories of this dynamical system.



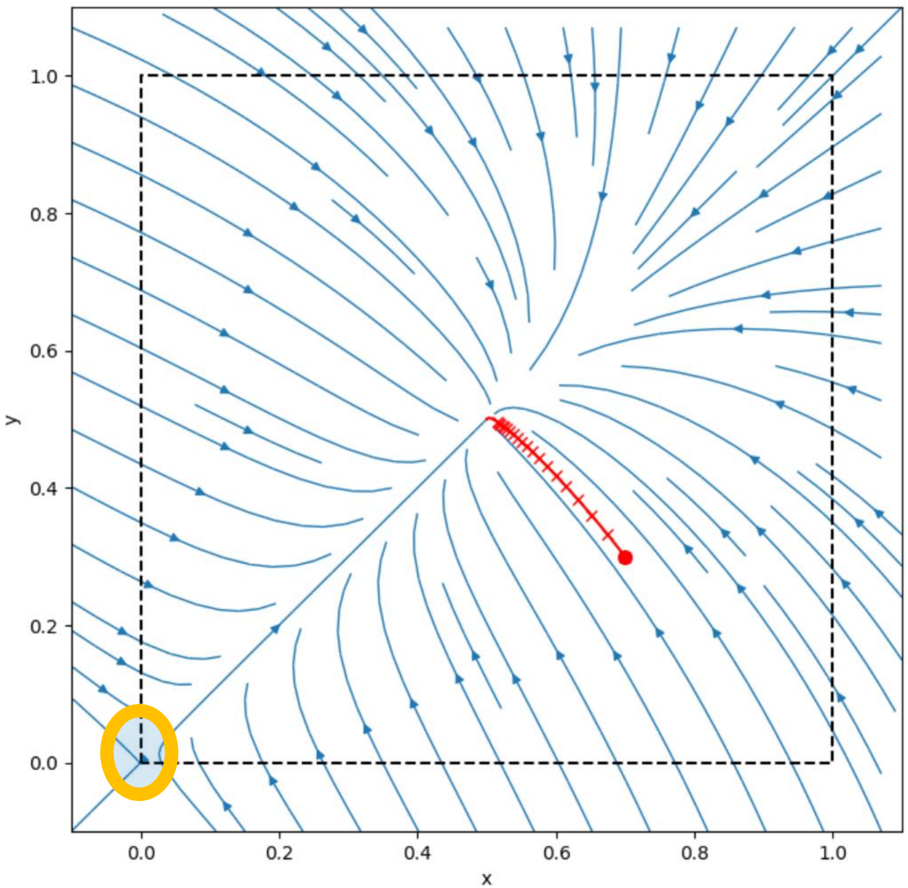
The setup



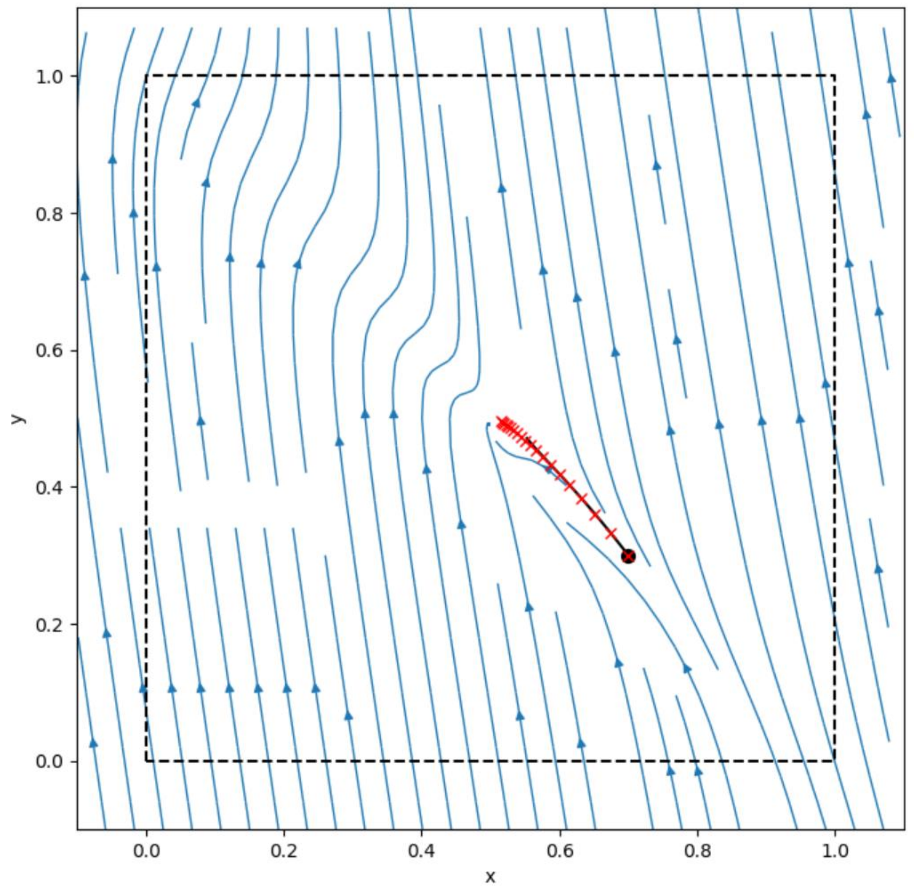
- The true dynamics f is unknown
- **What we observe:**
  Noisy measurements of the vector field on 20 points from a single trajectory starting from [0.7;0.3]
- **Goal:**
  - Learn a polynomial vector field p (of degree 2 or 3) that best agrees with the observed trajectory
  - Incorporate side information to generalize better to unobserved trajectories

Learning p of degree 2

The true dynamics f (unknown)



Least squares solution



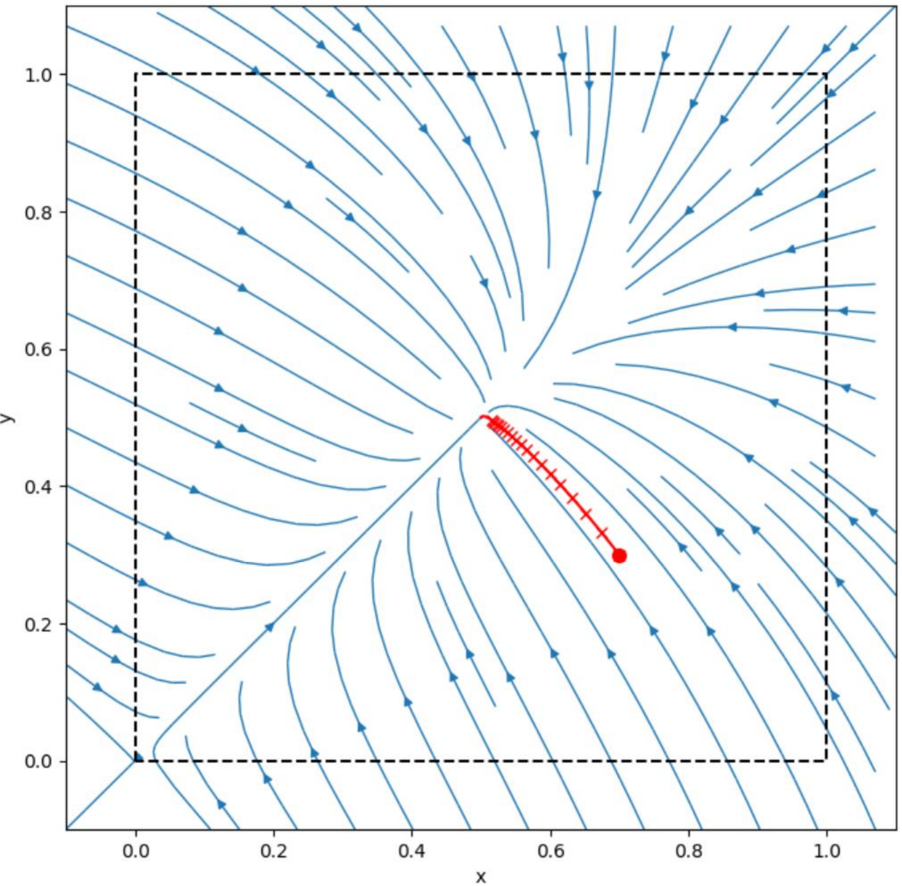
- Good performance on the observed trajectory. Horrible elsewhere.
- What side information can you think of in the context of this problem?

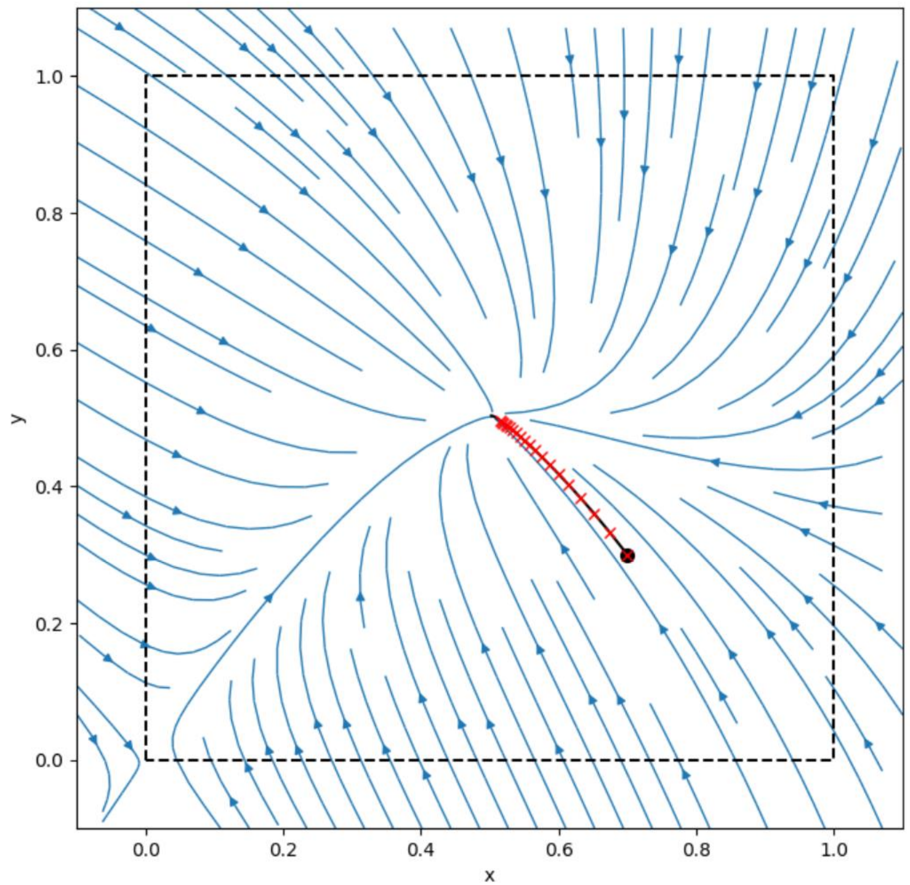
$\boldsymbol{f(0) = 0}$!!

Learning p of degree 2

The true dynamics f (unknown)

Least squares solution subject to $p(0) = 0$



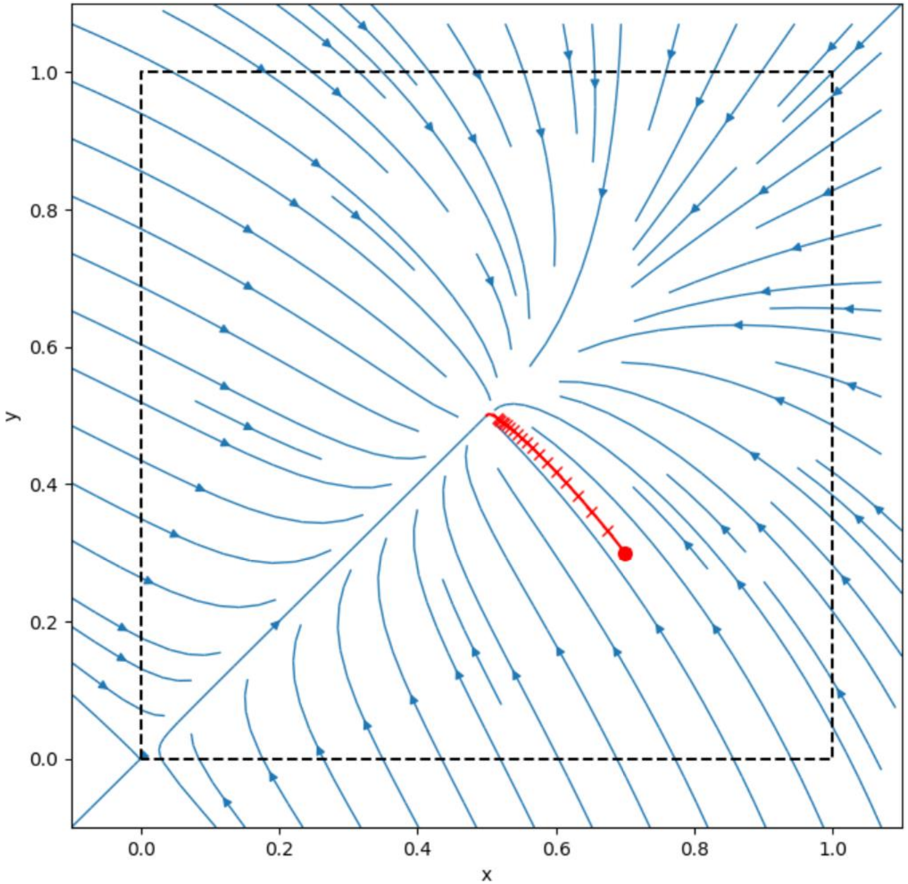


- Already gets the qualitative behavior on unobserved state space correctly!

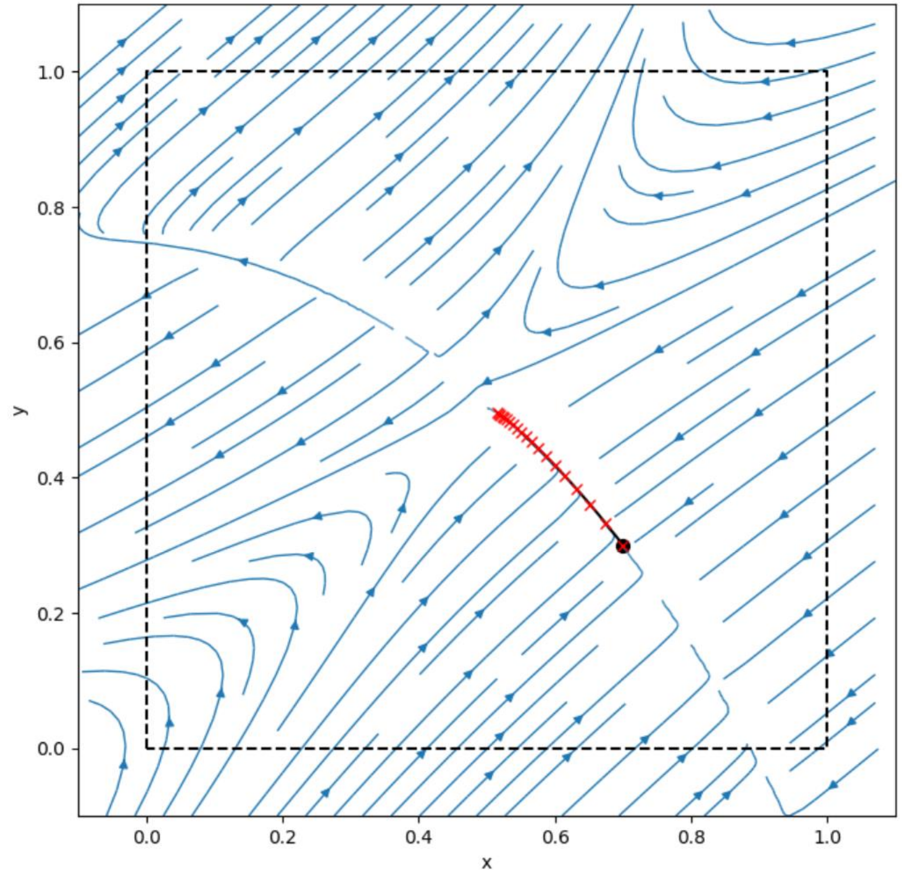
Maybe we are helping ourselves by taking p to have degree 2 (same as f)? Let's try higher degree p to assume a richer model than the truth.

Learning p of degree 3

The true dynamics f (unknown)



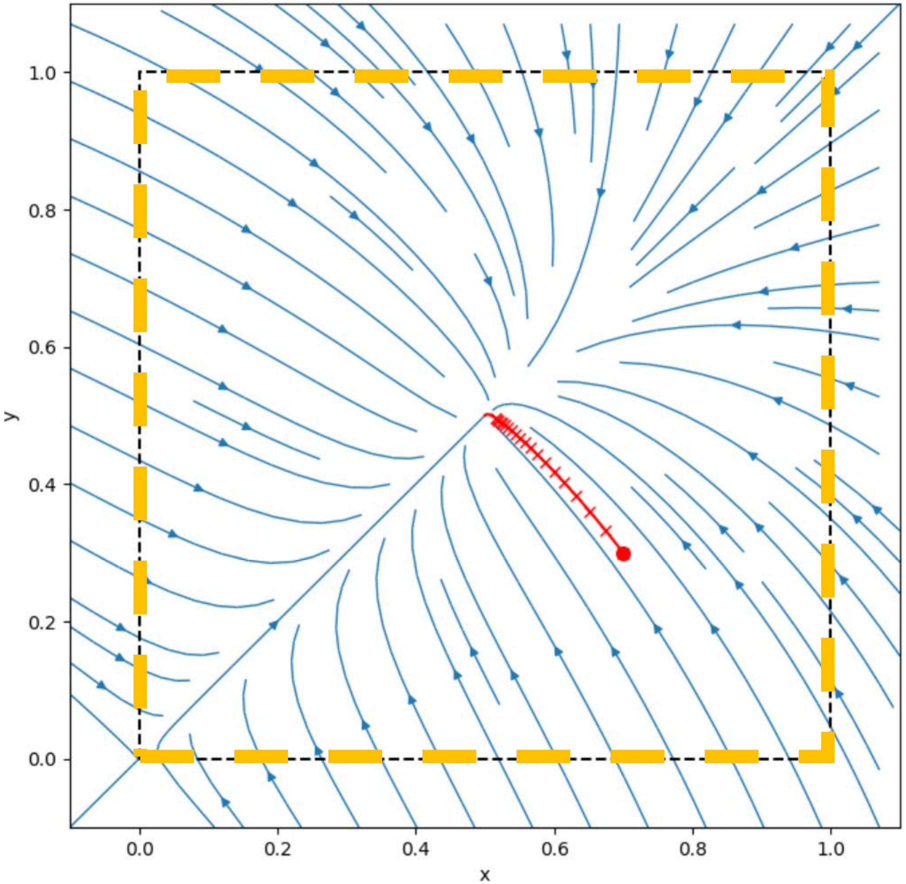
Least squares solution



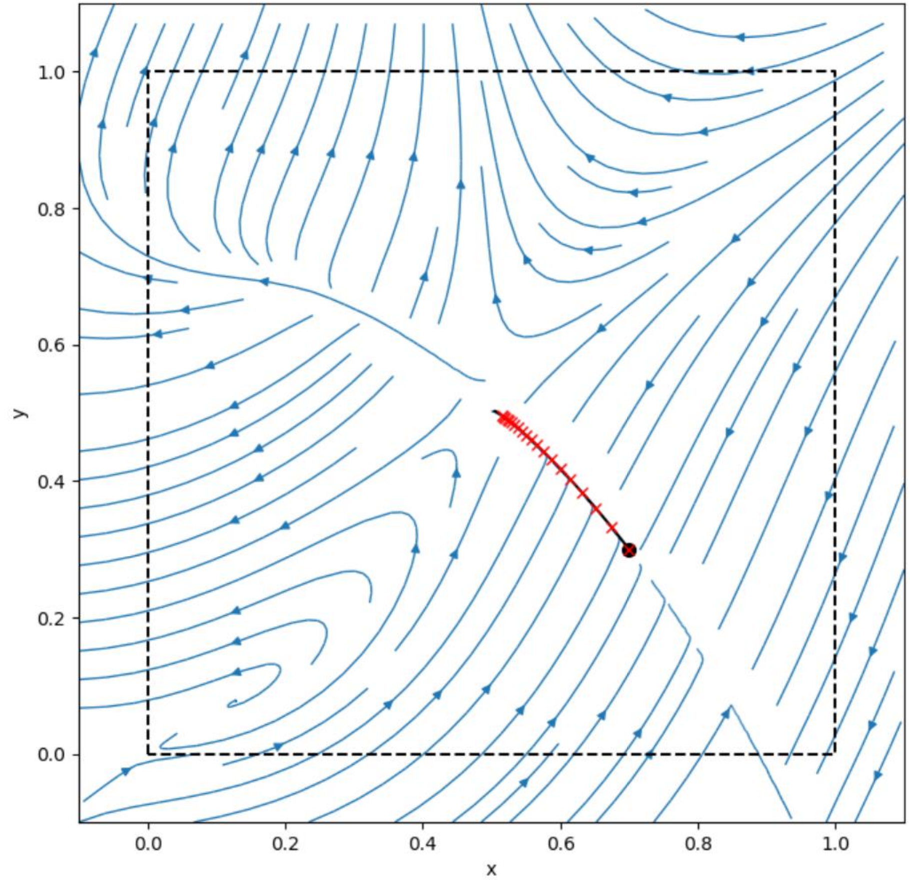
- Good performance on the observed trajectory. Again horrible elsewhere.

Learning p of degree 3

The true dynamics f (unknown)



Least squares solution subject to $p(0) = 0$



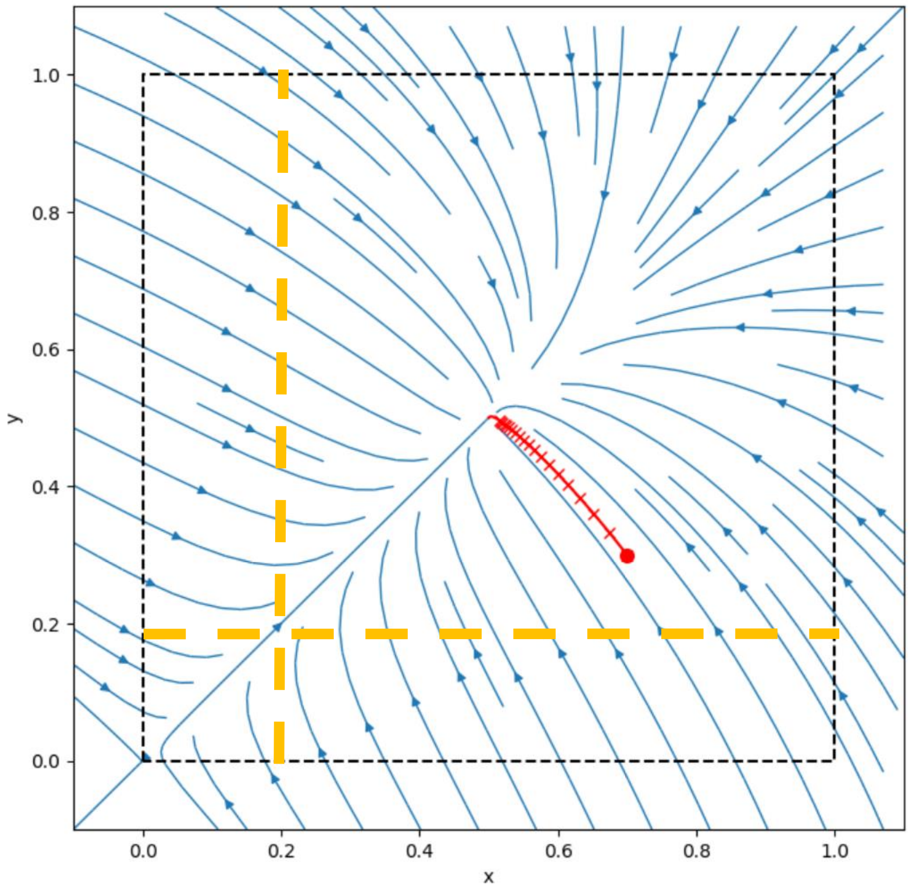
- Still pretty bad. What other side information can you think of?

Fraction of infected individuals cannot go negative or more than one!

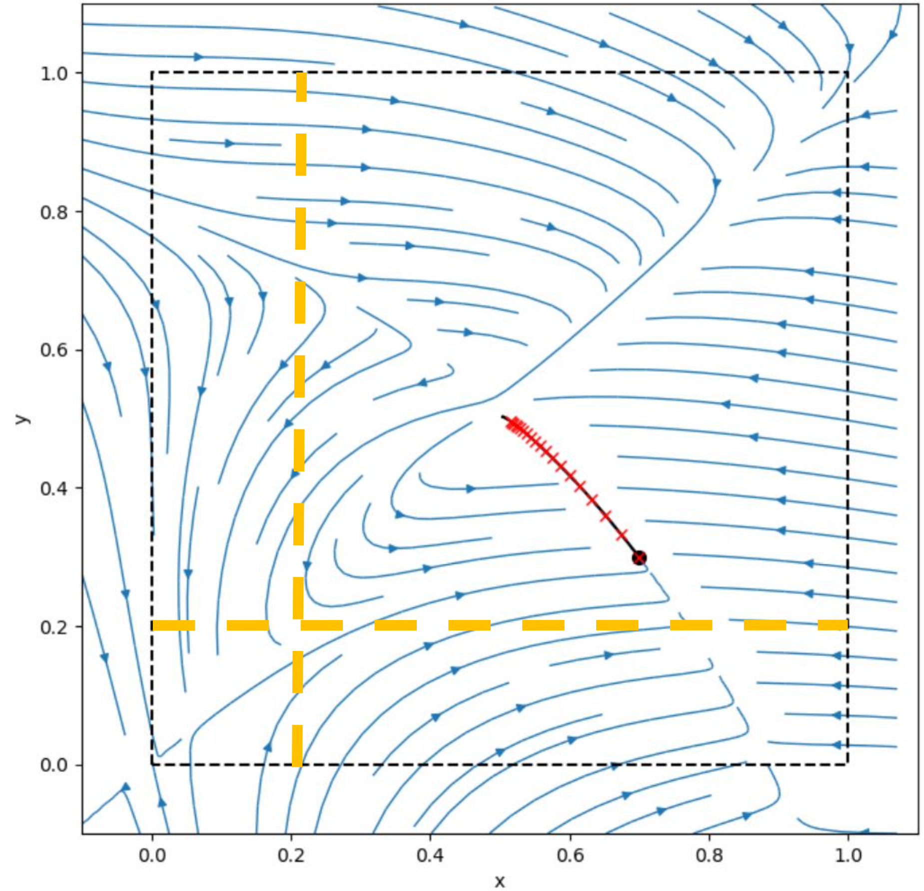
The unit square must be an invariant set!!

Learning p of degree 3

The true dynamics f (unknown)



Least squares solution subject to $p(0) = 0$, unit square invariant

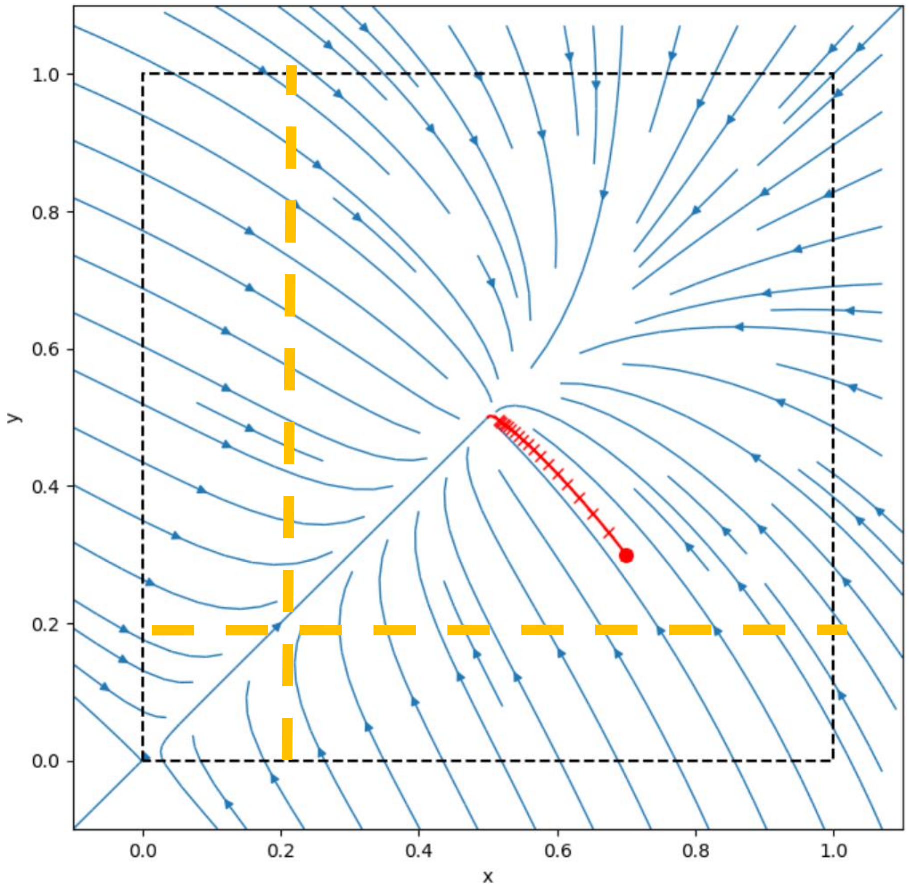


- Better, but not perfect. What other side information can you think of?

More infected females should imply higher infection rate for males! (and vice versa)

Side information: directional monotonicity

The true dynamics f (unknown)



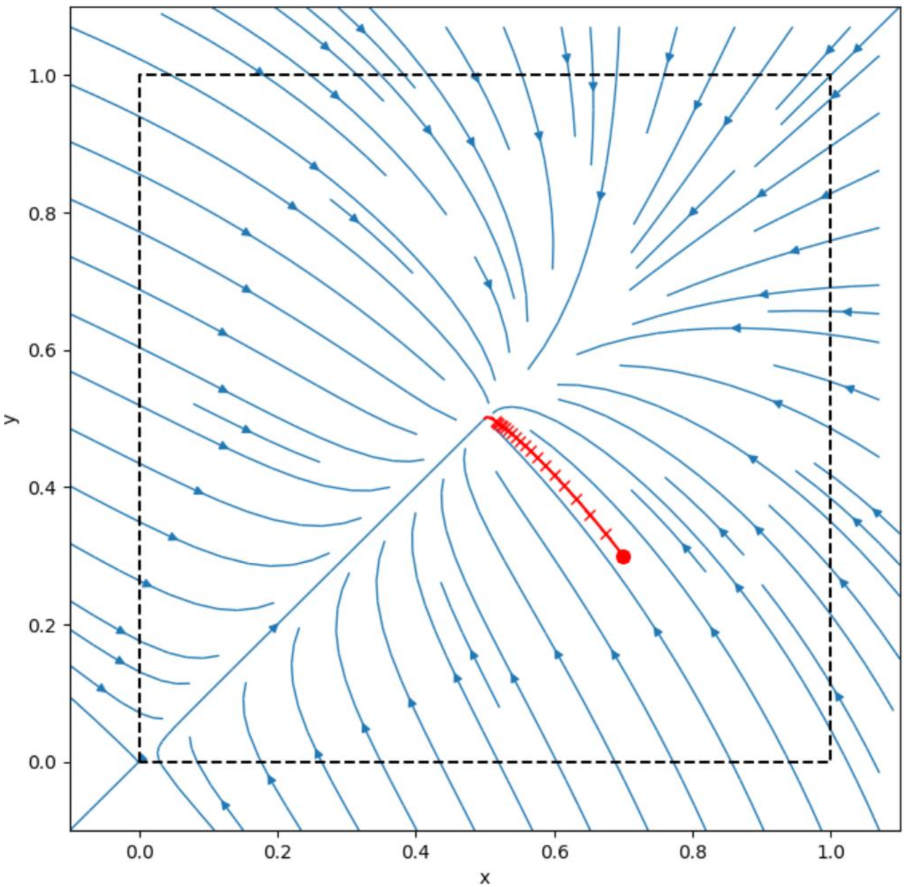
$$\frac{\partial f_1(x,y)}{\partial y} \geq 0, \forall (x,y) \in [0,1]^2$$

$$\frac{\partial f_2(x,y)}{\partial x} \geq 0, \forall (x,y) \in [0,1]^2$$

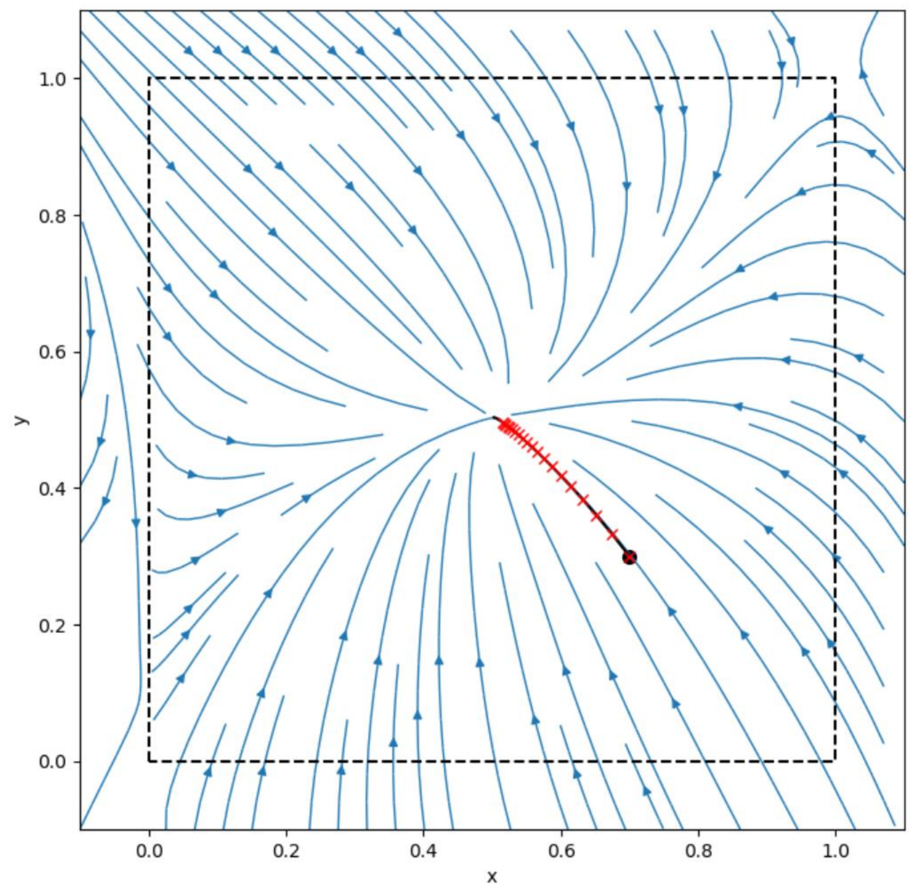
We want p to satisfy the same constraints!

Learning p of degree 3

The true dynamics f (unknown)



Least squares solution subject to $p(0) = 0$, unit square invariant, directional monotonicity

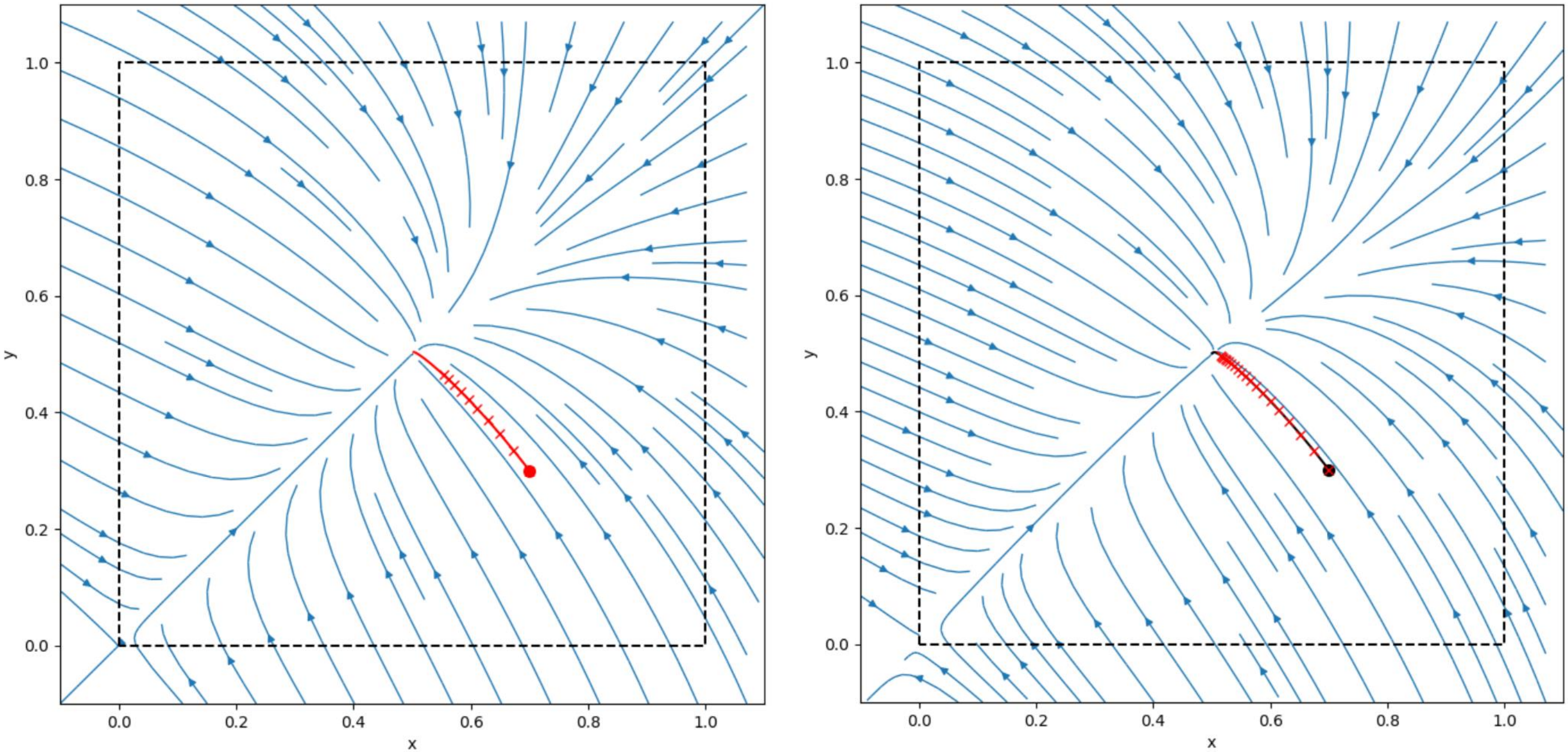


- Now we are getting the qualitative behavior correct everywhere!

Let's learn p of degree 2 again just for fun

The true dynamics f (unknown)

Least squares solution subject to $p(0) = 0$, unit square invariant, directional monotonicity



- p is pretty much dead on everywhere even though it was trained on a single trajectory!

The SDP that is being solved in the background

$$\min \sum_{i=1}^{20} \left(p(x^i, y^i) - \hat{f}(x^i, y^i) \right)^2$$

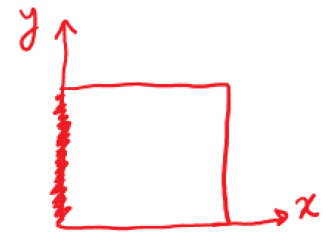
$p = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}, \deg(p) \leq 3$

$\sigma_0, \sigma_1, \deg(\sigma_i) \leq 2$

$\hat{\sigma}_0, \hat{\sigma}_1, \hat{\sigma}_2$

$\deg(\hat{\sigma}_0) \leq 2$

s.t. $p_1(0,0) = 0, \quad p_2(0,0) = 0$



$$p_1(0,y) = y\,\sigma_0(y) + (1-y)\,\sigma_1(y)$$

σ_0, σ_1 SOS

$\Rightarrow [x = 0, 0 \leq y \leq 1 \Rightarrow \dot{x} \geq 0]$

(+ three similar constraints)

$$\frac{\partial p_1}{\partial y}(x,y) = \hat{\sigma}_0(x,y) + \hat{\sigma}_1 (1-x)x + \hat{\sigma}_2 (1-y)y$$

$\hat{\sigma}_0$ SOS, $\hat{\sigma}_1 \geq 0, \hat{\sigma}_2 \geq 0$

$\Rightarrow \left[0 \leq x \leq 1, 0 \leq y \leq 1 \Rightarrow \frac{\partial p_1}{\partial y}(x,y) \geq 0\right]$

(similarly for $\frac{\partial p_2}{\partial x}(x,y)$)

Output of SDP solver:

p1=0.2681*x^3 - 0.0361*x^2*y - 0.095*x*y^2 + 0.1409*y^3 - 0.4399*x^2 + 0.0956*x*y - 0.0805*y^2 + 0.1232*x + 0.0201*y

p2=0.1188*x^3 + 0.2606*x^2*y + 0.2070*x*y^2 + 0.0005*y^3 - 0.3037*x^2 - 0.4809*x*y -0.099*y^2+ 0.2794*x+0.01689*y

Existence of constrained polynomial dynamics close to f

Thm [AAA, El Khadir]. For any continuously differentiable vector field $f: \mathbb{R}^n \to \mathbb{R}^n$, any $T > 0, \epsilon > 0$, and any compact set $\Omega \subseteq \mathbb{R}^n$,

there exists a polynomial vector field $p: \mathbb{R}^n \to \mathbb{R}^n$ such that

1) **trajectories of $\boldsymbol{f}$ and $\boldsymbol{p}$ starting from any initial conditions $\boldsymbol{x_0} \in \boldsymbol{\Omega}$ remain within $\boldsymbol{\epsilon}$ for all time $\boldsymbol{t} \in [\mathbf{0}, \boldsymbol{T}]$** (as long as they stay in Ω),

2) **$\boldsymbol{p}$ satisfies any combination of the following constraints if $\boldsymbol{f}$ does:**

a. equilibria at a given finite set of points ($p(v_i) = 0$),
b. invariance of a basic semialgebraic set $B = \{x \in \mathbb{R}^n | g_i(x) \geq 0\}$, where each g_i is concave (assumption can be relaxed),
c. directional monotonicity on a compact set ($\frac{\partial p_i(x)}{\partial x_j} \geq 0, \forall x \in C$),
d. nonnegativity on a compact set ($p_i(x) \geq 0, \forall x \in D$).

Moreover, all such properties of p come with an **SOS certificate**.

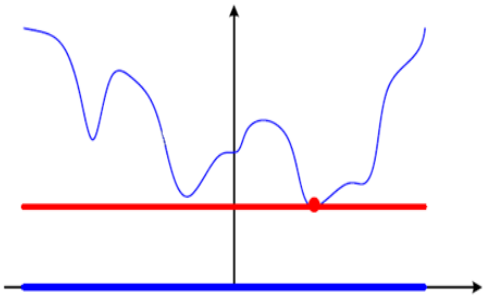
Proof ideas: Gronwall's inequality, constrained approximation theory, Putinar's Psatz

Recap

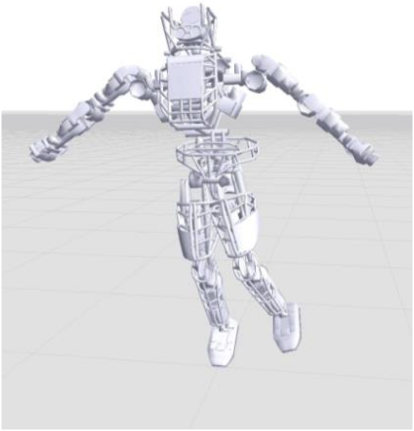
Is $p(x) \geq 0$ on $\{g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$?

Automated proofs via SOS/SDSOS/SOS!

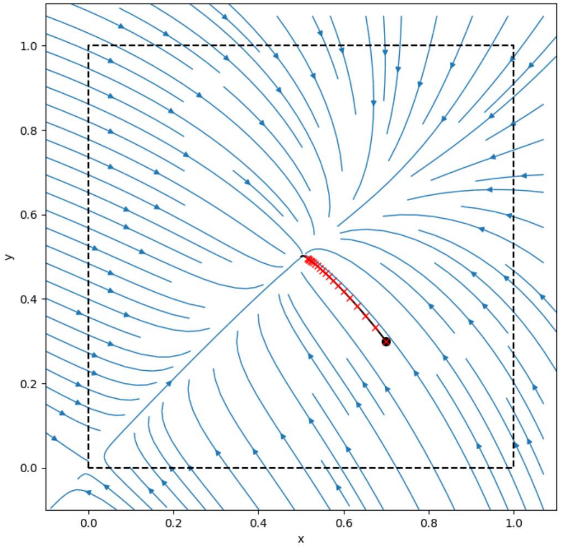
Many applications!



Optimization



Control



Learning

Want to know more? http://aaa.princeton.edu