Positively invariant set estimation: references, exercises, questions and answers

Didier Henrion[1,2]

Online learning weeks of the POEMA network

Lecture of June 3, 2020

1 References

The lecture follows closely [8], where all the proofs are detailed, in the case of controlled dynamical systems, both in continuous-time and discrete-time. Many of the technical arguments are adaptations of ideas proposed previously in [6] for approximating the region of attraction of controlled ordinary differential equations. In turn, the results in [6] can be seen as an adaptation to dynamical systems of the results of [7] dealing with approximation of the volume (and other moments) of a semi-algebraic sets. For further developments see [14].

The technical background on push-forward measures, Koopman and Frobenius-Perron operators for dynamical systems is covered in [12]. In particular, the example of push-forward measure for the logistic map is described in [12, Section 1.2].

The moment-SOS hierarchy was applied in [5] for computing (moments of) invariant measures (fixed points of the Frobenius-Perron operator) of one-dimensional dynamical systems, in particular for the logistic map. See [10, 13] for a broader perspective.

2 Exercises

2.1 Exercise 2.1

2.1.1 Statement

Consider the logistic map

$$f(x) = 4x(1-x)$$

[1]CNRS; LAAS; Université de Toulouse, 7 avenue du colonel Roche, F-31400 Toulouse, France.

[2]Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, CZ-16626 Prague, Czechia.

on

$$X := [0,1].$$

a. Given $\mu(dx) = m_0(x)dx$, derive analytically $f_\#\mu$.

b. Given $\mu(dx) = I_{[0,1]}(x)dx$ compute $f_\#\mu$ and $f \circ f_\#\mu$.

c. Prove that $\mu(dx) = dx/(\pi\sqrt{x(1-x)})$ is invariant.

d. Prove that $\mu(dx) = \delta_{3/4}(dx)$ is invariant.

2.1.2 Solution

a. By definition

$$f_\#(A) = \mu \circ f^{-1}(A)$$

where $f^{-1}(A) := \{x \in X : f(x) \in A\}$. If $A = [0,x] \subset X$ we can check easily that

$$f^{-1}([0,x]) = [0, \frac{1}{2} - \frac{1}{2}\sqrt{1-x}] \cup [\frac{1}{2} + \frac{1}{2}\sqrt{1-x}, 1].$$

Therefore

$$f_\#\mu([0,x]) = \int_0^x f_\#\mu(dy) = \int_{f^{-1}([0,x])} \mu(dy)$$

and

$$\int_0^x m_1(y)dy = \int_{f^{-1}([0,x])} m_0(y)dy$$

if μ resp. $f_\#\mu$ have density m_0 resp. m_1 with respect to the Lebesgue measure. Differentiating with respect to x yields

$$\begin{aligned} m_1(x) &= \frac{d}{dx}\int_{f^{-1}([0,x])} m_0(y)dy \\ &= \frac{d}{dx}\int_0^{\frac{1}{2}-\frac{1}{2}\sqrt{1-x}} m_0(y)dy + \frac{d}{dx}\int_{\frac{1}{2}+\frac{1}{2}\sqrt{1-x}}^1 m_0(y)dy \\ &= \tfrac{1}{4\sqrt{1-x}}\left(m_0(\tfrac{1}{2} - \tfrac{1}{2}\sqrt{1-x}) + m_0(\tfrac{1}{2} + \tfrac{1}{2}\sqrt{1-x})\right). \end{aligned}$$

b. From the above formula we get

$$f_\# I_{[0,1]}(x)dx = \frac{dx}{2\sqrt{1-x}}$$

and

$$(f \circ f)_\# I_{[0,1]}(x)dx = f_\# \frac{dx}{2\sqrt{1-x}} = \left(\frac{1}{\sqrt{1+\sqrt{1-x}}} + \frac{1}{\sqrt{1-\sqrt{1-x}}}\right)\frac{\sqrt{2}dx}{8\sqrt{1-x}}.$$

c. This follows readily from the above formula by replacing m_0 and m_1 with $(\pi\sqrt{1-x})^{-1}$. Probability measures μ satisfying $f_\#\mu = \mu$ are called invariant measures.

d. The identity $f_\# \mu = \mu$ also writes $\langle v, f_\# \mu \rangle = \langle v(f), \mu \rangle = \langle v, \mu \rangle$ for all test functions $v \in C(X)$. In particular if $\mu = \delta_x$ this yields $v(f(x)) = v(x)$. It can be checked readily that $x = 3/4$ satisfies $f(x) = x$ and hence that $\delta_{3/4}$ is an invariant measure.

Note also that every finite length orbit $(x_1, x_2 = f(x_1), \ldots, x_N = f(x_{N-1}), x_1 = f(x_N))$ in X yields an invariant measure $\mu = \frac{1}{N}\sum_{k=1} \delta_{x_k}$ since $\langle v(f), \mu \rangle = \frac{1}{N}\sum_{k=1}^{N} v(f(x_k)) = \frac{1}{N}(\sum_{k=1}^{N-1} v(x_{k+1}) + v(x_1)) = \langle v, \mu \rangle$ for all $v \in C(X)$.

2.2 Exercise 2.2

2.2.1 Statement

Consider the LP

$$\begin{array}{rcll} p^* & = & \sup & \langle 1, \mu_0 \rangle \\ & & \text{s.t.} & \mu = \mu_0 + \alpha f_\# \mu \\ & & & \mu_0 + \hat{\mu}_0 = \lambda_X \end{array}$$

where λ_X is the Lebesgue measure on X and the optimization variables are μ, μ_0, $\hat{\mu}_0$ all in $C(X)'_+$.

Prove that the supremum is attained by $\mu_0^* = \lambda_{X_I}$ and hence $p^* = \mathrm{vol}\, X_I$.

2.2.2 Solution

For any initial measure μ_0 with $\mu_0 \leq \lambda_X$, i.e. with density less than one on X, and with $\mathrm{spt}\, \mu_0 \subset X_I$, there exists an occupation measure μ solving the Liouville equation $\mu = \mu_0 + \alpha f_\# \mu$. One such measure is $\mu_0 = \lambda_{X_I}$. Together with the slack measure $\hat{\mu}_0 = \lambda_{X \setminus X_I}$, the triplet $(\mu, \mu_0, \hat{\mu}_0)$ is feasible for the LP, and hence $p^* \geq \langle 1, \mu_0 \rangle = \mathrm{vol}\, X_I$.

We have seen also that for any pair of measures (μ, μ_0) satisfying the Liouville equation with $\mathrm{spt}\, \mu$ and $\mathrm{spt}\, \mu_0$ in X, it holds $\mathrm{spt}\, \mu_0 \subset X_I$. This proves that $p^* \leq \mathrm{vol}\, X_I$.

2.3 Exercise 2.3

2.3.1 Statement

Derive the dual LP

$$\begin{array}{rcll} d^* & = & \inf & \langle w, \lambda_X \rangle \\ & & \text{s.t.} & (v - \alpha v \circ f,\ w - v - 1,\ w) \in C(X)^3_+. \end{array}$$

by convex duality. Prove that there is no duality gap.

2.3.2 Solution

Let us start with the primal in minimization form

$$\begin{array}{rcll} -p^* & = & \inf & \langle -1, \mu_0 \rangle \\ & & \text{s.t.} & \mu = \mu_0 + \alpha f_\# \mu \\ & & & \mu_0 + \hat{\mu}_0 = \lambda_X \\ & & & (\mu, \mu_0, \hat{\mu}_0) \in (C(X)'_+)^3 \end{array}$$

and construct the Lagrangian

$$\ell(\mu, \mu_0, \hat{\mu}_0, v, w) := \langle -1, \mu_0 \rangle + \langle v, \mu - \mu_0 - \alpha f_\# \mu \rangle + \langle w, \mu_0 + \hat{\mu}_0 - \lambda_X \rangle$$

where the dual variable $v \in C(X)$ corresponds to the Liouville equation, and the dual variable $w \in C(X)$ corresponds to the Lebesgue domination equation. Rearrange the Lagrangian

$$\ell(\mu, \mu_0, \hat{\mu}_0, v, w) = \langle -w, \lambda_X \rangle + \langle v - \alpha v \circ f, \mu \rangle + \langle -1 - v + w, \mu_0 \rangle + \langle w, \hat{\mu}_0 \rangle$$

such that the dual Lagrange function can be expressed as

$$d(v, w) := \inf_{\mu, \mu_0, \hat{\mu}_0} \ell(\mu, \mu_0, \hat{\mu}_0, v, w) = \langle -w, \lambda_X \rangle$$

provided $(v - \alpha v \circ f, -1 - v + w, w) \in C(X)^3_+$. Maximization of the Lagrange function yields the dual LP

$$\begin{array}{rcll} -d^* & = & \sup & \langle w, -\lambda_X \rangle \\ & & \text{s.t.} & (v - \alpha v \circ f,\ w - v - 1,\ w) \in C(X)^3_+. \end{array}$$

To prove that there is no duality gap, we use [2, Theorem IV.7.2] and the fact that the feasible set of the primal LP is nonempty and bounded in the metric inducing the weak-star topology on measures. To see non-emptiness, notice that the vector of measures $(\mu, \mu_0, \hat{\mu}_0) = (0, 0, \lambda_X)$ is trivially feasible. To see the boundedness, it suffices to integrate the equality constraints of the primal LP on the whole domain X. This gives $\mu_0(X) + \hat{\mu}_0(X) = \lambda(X) < \infty$ and $\mu(X) = \mu_0(X)/(1 - \alpha) < \infty$ since $\alpha \in (0, 1)$ and all the measures are non-negative.

2.4 Exercise 2.4

2.4.1 Statement

By replacing $C(X)_+$ with $Q(X)_{r,r}$ we get a monotone converging sequence of upper bounds

$$p^*_r = d^*_r \geq p^*_{r+1} = d^*_{r+1} \geq p^*_\infty = d^*_\infty = \text{vol}\, X_I.$$

Prove it with the Stone-Weierstrass Theorem.

2.4.2 Solution

The absence of duality gap in the moment-SOS hierarchy, i.e. $p^*_r = d^*_r$ for all r, follows from the same primal boundedness arguments as in Exercise 2.3. These are upper bounds since the dual quadratic module $Q(X)'_{r,r}$ is an outer approximation of the moment cone $P(X)'_r$ and the objective function is maximized in the primal moment relaxation. Monotonicity of the sequence of bounds follows from the embedded structure of the truncated quadratic modules, i.e. $Q(X)_{r,r} \subset Q(X)_{r+1,r+1}$. Convergence of the bounds follows from the Stone-Weierstrass Theorem [1, A7.5] which asserts that a continuous function on a compact set $X \subset \mathbb{R}^n$ can be approximated arbitrarily well (in the strong topology of the supremum norm) by polynomials. In turn, these polynomials can be approximated arbitrarily well with SOS (and hence semidefinite programming) by Putinar's Positivstellensatz, provided the quadratic module is Archmedian, as already discussed many times in this course.

2.5 Exercise 2.5

2.5.1 Statement

In the dual we obtain a sequence of polynomials v_r, w_r in $\mathbb{R}[x]_r$ such that

$$X_{Ir} := \{x \in X : v_r(x) \geq 0\} \supset X_I$$

and

$$\lim_{r\to\infty} \mathrm{vol}(X_{Ir} \setminus X_I) = 0.$$

Prove it by showing that $w_r \to I_{X_I}$ in $\mathscr{L}_1(X)$.

2.5.2 Solution

We know from the constraints of the dual SOS problem that $w_r \geq I_{X_I}$. Since the sequence of values of the relaxations $\|w_r\|_{\mathscr{L}_1(X)} = \int_X w_r(x)dx$ converges from above to $\int_X I_{X_I}(x)dx =$ vol X_I, it follows that w_r converges in $\mathscr{L}_1(X)$ norm. We know that $X_{Ir} \supset X_I$ and since $w_r \geq I_{X_{I_r}} \geq I_{X_I}$ we also have $\{x \in X : w_r(x) \geq 1\} \supset X_{Ir} \supset X_I$. Since w_r converges to I_{X_I} in $\mathscr{L}_1(X)$ norm, it holds $\lambda(X_I) = \lim_{r\to\infty} \int_X w_r(x)dx \geq \lim_{r\to\infty} \int_X I_{X_{I_r}}(x)dx = \lim_{r\to\infty} \lambda(X_{Ir}) \geq \lim_{r\to\infty} \lambda(\cap_{k=1}^r X_{Ik}) = \lambda(\cap_{k=1}^\infty X_{Ik})$. Since $X_I \subset X_{Ir}$ for all r we have $\lim_{r\to\infty} \lambda(X_{Ir}) = \lambda(X_I)$ and $\lambda(\cap_{k=1}^\infty X_{Ik}) = \lambda(X_I)$.

3 Questions and answers

Q. Are the eigenvectors of Frobenius-Perron operator known ?

A. They can be approximated numerically with the moment-SOS hierarchy, see [10, Section 8] and references therein. The study of the eigenstructure of the Frobenius-Perron operator (and its adjoint the Koopman operator) is still a vast subject of research. For references on the numerics of the Koopman operator in dynamical systems and optimization see e.g. [3] and more recently [11] in the context of data-driven methods.

Q. The primal and dual convex problems depend on the discount rate α. This cannot be optimized, because it multiplies the unknown measures/continuous functions. Is α to be chosen arbitrarily? If so, does the feasibility of these convex problems depend on the particular value of α?

A. The theoretical convergence results do not depend on the discount rate α. The numerics of solving the optimization problems in the moment-SOS hierarchy will indeed depend on the discount rate, but in a way which is difficult to understand. Generally, a value of α close to 1 gives satisfactory results.

Q. Anything known about the convergence rate of the bounds?

A. Our MPI set approximation method can be viewed as an extension of the volume approximation method of [7]. In this previous work, the moment-SOS hierarchy was used to approximate as closely as desired the volume (and all the other moments) of a given semi-algebraic set. Convergence rates for this volume approximation were studied in [9]. They are quite pessimistic, of the order of $\log\log 1/r$ where r is the relaxation order. In practice, we observe a sublinear rate of convergence. The convergence results of [9] may be extended to the MPI set approximation, but the essential difference is that in [7] the set to be approximated is given. The geometry of the set to be approximated plays a key role in the convergence analysis of [9], and we know that the geometry of the MPI set can be intricate (e.g. with a fractal boundary for the Julia sets).

Q. Can the approach be applied to approximate other "type" of sets/graph ? (partially in the previous question answer)

A. An interesting application domain could be shape optimization. For example, Newton's problem of finding the 3D convex body of minimal resistance can be formulated as a linear optimization problem where the unknown is the surface area measure of the convex body [4]. In principle the moment-SOS hierarchy can be applied to solve numerically this problem. The graph of the body can then be recovered approximately with the Christoffel-Darboux kernel techniques of [15].

Q. Are there any difficulties in extending the analysis to continuous-time systems, especially in terms of convergence results?

A. The results of [8], on which this course is based, are also available in continuous-time. See also [6] for approximations of the region of attraction of continuous-time ordinary differential equations.

Q. What is the (dis)advantages of Lasserre hierarchy approach in finding the approximations (or the volume) of an invariant set compared to other methods in theoretical and practical respects?

A. We are missing comprehensive numerical comparisons with alternative techniques for set estimation. For recent references on volume computation, see [16] and references therein.

Q. Can another outer measure (not Lebesgue; e.g., Hausdorff) be used to compute the volume of the set X_I so we can design iterative methods converging in this other measure?

A. Indeed, one can use another outer measure than λ_X in the primal-dual LP, as soon as we can compute accurately all its moments. These moments must be available since they enter explicitly in the coefficients of the moment-SOS hierarchy.

References

[1] R. Ash. Real analysis and probability. Academic Press, 1972.

[2] A. Barvinok. A course in convexity. AMS, 2002.

[3] M. Budisić, R. Mohr, I. Mezić. Applied Koopmanism. Chaos: An Interdisciplinary Journal of Nonlinear Science, 22(4):047510, 2012.

[4] G. Carlier, T. Lachand-Robert. Convex bodies of optimal shape. Journal of Convex Analysis, 10(1):265–27, 2003.

[5] D. Henrion. Semidefinite characterisation of invariant measures for one-dimensional discrete dynamical systems. Kybernetika, 48(6):1089-1099, 2012.

[6] D. Henrion, M. Korda. Convex computation of the region of attraction of polynomial control systems. IEEE Transactions on Automatic Control, 59(2):297-312, 2014.

[7] D. Henrion, J. B. Lasserre, C. Savorgnan. Approximate volume and integration for basic semialgebraic sets. SIAM Review, 51(4):722-743, 2009.

[8] M. Korda, D. Henrion, C. N. Jones. Convex computation of the maximum controlled invariant set for polynomial control systems. SIAM Journal on Control and Optimization, 52(5):2944-2969, 2014.

[9] M. Korda, D. Henrion. Convergence rates of moment-sum-of-squares hierarchies for volume approximation of semialgebraic sets. Optimization Letters, 12(3):453-442, 2018.

[10] M. Korda, D. Henrion, I. Mezić. Convex computation of extremal invariant measures of nonlinear dynamical systems and Markov processes. `arXiv:1807.08956`, 2018.

[11] M. Korda, M. Putinar, I. Mezić. Data-driven spectral analysis of the Koopman operator. Applied and Computational Harmonic Analysis 48(2):599-629, 2020.

[12] A. Lasota, M. C. Mackey. Probabilistic properties of deterministic systems. Cambridge University Press, 1985.

[13] V. Magron, M. Forets, D. Henrion. Semidefinite approximation of invariant measures for polynomial systems. Discrete and Continuous Dynamical Systems Series B, 24(12):6745-6770, 2019.

[14] V. Magron, P.-L. Garoche, D. Henrion, X. Thirioux. Semidefinite approximations of reachability sets for discrete-time polynomial systems. SIAM Journal on Control and Optimization, 57(4):2799–2820, 2019.

[15] S. Marx, E. Pauwels, T. Weisser, D. Henrion, J. B. Lasserre. Tractable semi-algebraic approximation using Christoffel-Darboux kernel. `arXiv:1904.01833`, 2019.

[16] M. Tacchi, T. Weisser, J. B. Lasserre, D. Henrion. Exploiting sparsity for semi-algebraic set volume computation. `arXiv:1902.02976`, 2019.