Path Planning for Mobile Inference of Spatiotemporally Evolving Systems

Joshua Whitman, Girish Chowdhary
Coordinated Science Laboratory (CSL)
University of Illinois
Urbana, IL 61801
{jewhitm2,girishc}@illinois.edu

Abstract

We present a new method for planning trajectories of moving agents with the goal of deriving the state of a spatiotemporally evolving dynamical systems. This method is based on the spectral analysis of the linear dynamical layer of an Evolving Gaussian Processes (E-GP) model. First, we present a new algorithm for clustering elements of the system into separate invariant subspaces. Next, we discuss the relative importance of several factors in generating and choosing trajectories that help a moving Bayesian state estimator converge to an accurate state estimate quickest. Lastly, we provide preliminary results for these new methods using synthetic and real-world data sets.

1 Introduction

One of the most difficult and important problems in bringing machine learning to physical domains is modeling and tracking large-scale stochastic phenomena with both spatial and temporal (spatiotemporal) evolution [1]. Examples of such phenomena include temperature variation, CO_2 flux over large areas, extreme weather events [5] like wildfires, pedestrian traffic patterns, soft robotics, and fluid dynamics. The latter is governed by the Navier-Stokes partial differential equations. Numerical solutions to these equations are highly resource-intensive, making them poorly suited for online robotic applications, such as autonomous aerial, ground, or water vehicles.

In contrast to first-principles approaches, data-driven models of spatiotemporally evolving phenomena have been gaining more attention in the machine learning and statistics communities [4]. In 2017, a highly efficient approach was presented called "Evolving Gaussian Processes" (E-GPs) which aimed to replace costly numerical simulations for design and autonomy purposes [10], by using a machine learning model that could generalize across similar physical situations. It is the goal of this paper to extend that work from offline simulation to online inference via measurements from mobile agents.

1.1 Contribution

We believe we have made significant advances in choosing optimal paths for a mobile agent to take in a domain in order to infer the state of a system that is changing in both space and time. The algorithms presented in this paper provide promising results in both simple and complex synthetic systems as well as in the complex physics governed the Navier-Stokes partial differential equation (PDEs). We build upon very recent results published in [10] and [7], using spectral analysis to illuminate features of the dynamic model which are useful for determining the best paths for a mobile agent to follow.

The essential idea behind Evolving Gaussian Processes is layer a linear dynamic transition model over the weights of a kernel-based model, such as a Gaussian Process with a Radial Basis Function kernel.

Workshop on Modeling and Decision-Making in the Spatiotemporal Domain, 32nd Conference on Neural Information Processing Systems (NIPS 2018), Montréal, Canada.

This separation of spatial and temporal components is what makes the learning problem tractable, even while modeling very complex systems. The evolution of large function spaces governed by complex PDEs is *transformed* into a evolution of "weight vectors" in a relatively smaller Hilbert space, which is spatially encoded with the original domain via the kernels. By analyzing the structure of the linear transition model in the weight space, we can obtain provides key physical insights into the system, which enable the design of observers and controllers [7, 6], as well as the path-planning algorithms presented here.

We present three main algorithmic contributions in this paper. First, we present a method, inspired by k-means clustering, for identifying areas of the domain which correspond to invariant subspaces in the linear transition matrix. Secondly, we present a method for identifying which path amongst a library of paths is most likely to provide the most useful data for a moving agent in the system. Thirdly, we present a method for generating useful paths based on the first clustering algorithm. Lastly, we demonstrate our methods on both some simple synthetic examples as well as on complex CFD data.

2 Background

2.1 Kernel Observers & Evolving Gaussian Processes

As presented in [7], let $y \in \mathbb{R}^N$ be the measurements of the function available from N sensors, $\mathcal{A} : \mathcal{H} \to \mathcal{H}$ be a linear transition operator in the Reproducing Kernel Hilbert Space $\mathcal{H}$ generated by the kernel k, and $\mathcal{K} : \mathcal{H} \to \mathbb{R}^N$ be a linear measurement operator. The model for our system is:

$$f_{\tau+1} = \mathcal{A}f_\tau + \eta_\tau, \quad y_\tau = \mathcal{K}f_\tau + \zeta_\tau, \tag{1}$$

where η_τ and ζ_τ each represent noise in their respective domains. We work with an approximate feature map $\widehat{\psi}(x) := [\,\widehat{\psi}_1(x) \cdots \widehat{\psi}_M(x)\,]$ to an approximate feature space $\widehat{\mathcal{H}}$. This is defined by the dictionary of atoms $\mathcal{F}^{\mathcal{C}} = \{\psi(c_1), \cdots, \psi(c_M)\}$, $\psi(c_i) \in \mathcal{H}$, $\widehat{\psi}_i(x) := k(x, c_i)$, (where $\mathcal{C} = \{c_1, \ldots, c_M\}$, $c_i \in \Omega$ compact), and thus $\widehat{\mathcal{H}} = \text{span}\mathcal{F}^{\mathcal{C}} \subseteq \mathcal{H}$. The finite-dimensional evolution equations approximating (1) in approximate dual form are

$$w_{\tau+1} = \widehat{A}w_\tau + \eta_\tau, \quad y_\tau = Kw_\tau + \zeta_\tau, \tag{2}$$

where we have a transition matrix $\widehat{A} \in \mathbb{R}^{M\times M}$, the kernel matrix $K = [\,\widehat{\psi}_1(x) \cdots \widehat{\psi}_M(x)\,] \in \mathbb{R}^{N\times M}$, the weight vectors $w_\tau \in \mathbb{R}^M$, and we abuse notation slightly to let η_τ and ζ_τ be noise again.

In this paper, it will be assumed that an E-GP model has already been learned from historical data using the 3 (or 4) step process summarized in [10], either specific to the system or from a family of similar systems. The purpose of this paper is to choose paths that allow a moving agent to rapidly infer the *state* of the system from single measurements.

3 Spectral Analysis of E-GP Model and Resulting Algorithms

An invariant subspace of a linear map $\widehat{A}$ on a vector space is a subspace W of $\mathbb{R}^M$ that is preserved by $\widehat{A}$, i.e. $\widehat{A}(W) \subseteq W$, and thus information contained in an invariant subspace never leaves it. Trivial examples include $\{0\}$ and $\mathbb{R}^M$, but we are interested dividing the domain into a set of regions by subspaces. We can do this because the weights in an E-GP are *spatially encoded*, i.e. $\{w\}_i \leftrightarrow c_i$, unlike in a neural network. We hypothesize that the c_i associated with the invariant subspaces of a spatiotemporally evolving system are generally associated with spatial regions in the domain. This hypothesis makes sense: physically, we know the *principle of locality* states that an object is only directly influenced by its immediate surroundings [2], and mathematically, thanks to the RBF kernel functions we expect nearby centers to be connected dynamically.

When a square matrix is perfectly formed, the Jordan form is block diagonal and gives us the invariant subspaces immediately. In reality, data-driven approximations of $\widehat{A}$ rarely give such easily interpretable properties. Hence the need for an algorithm with divides invariant subspaces.

3.1 k-Invariant Subspaces Clustering

The intuition behind our k-invariant subspaces clustering algorithm is to replace the Euclidean distance in the k-Means algorithm with a different metric of "nearness", namely one corresponding with

the dynamic connections. The $\widehat{A}$ matrix provides easy access to these: its rows $\widehat{A}_{i*}$ indicate which centers inform the i$^{\text{th}}$ value of $w_{\tau+1}$, and its columns $\widehat{A}_{*j}$ indicate which centers will be informed by the j$^{\text{th}}$ value of w_τ. We condition $\widehat{A}$ by first removing the eigenvectors whose eigenvalues are not on the unit circle, and then by controlling for the differing eigenmode frequencies, ending up with $\bar{A} = \bar{U}\bar{U}^\dagger$. Much like the pairwise squared deviations of points formulation of the k-means clustering problem, we can now write our problem as

$$\arg\max_S \sum_{i=1}^{k} \frac{1}{2|S_i|} \sum_{\substack{x_i \neq x_j \\ x_j, x_i \in S_i}} \bar{A}_{ij}^2$$

This problem can be solved with the following algorithm:

Algorithm 1 k-Invariant Subspaces Algorithm

while clusters have changed **do**
  for each center i **do**
    find cluster k which maximizes the score

$$\frac{1}{|S_k|+1}\left(\|\bar{A}_{i,S_k\setminus\{i\}}\|^2 + \|\bar{A}_{S_k\setminus\{i\},i}\|^2\right)$$

    reassign center i to cluster with highest score
  end for
end while
return clusters

3.2 Scoring Paths

We postulate that the value (with respect to a particular eigenvector) of taking a measurement at a point in the domain is proportional to the following factors. This scoring method allows us to decide the likeliest candidate amongst a large family of agent paths:

1. The spatial extent of the eigenvector v_i. This can be taken to equal the area under the curve of the magnitude of the complex function $v_i\hat{\Phi}(x)$, normalized by the its peak.

$$SE_i = \frac{1}{\sup_x v_i\hat{\Phi}(x)} \int_X |v_i\hat{\Phi}(x)|dx$$

2. The expected size of the measurement itself, which is roughly equal to $v_i\hat{\Phi}(x)$.
3. The ratio of the frequency with which that Koopman mode is visited to the frequency of its eigenvalue. This factor is included due to the aliasing effect
4. A discount factor equal to a decaying exponential of the number of times that eigenmode has been visited

3.3 Generating High-Scoring Random Paths

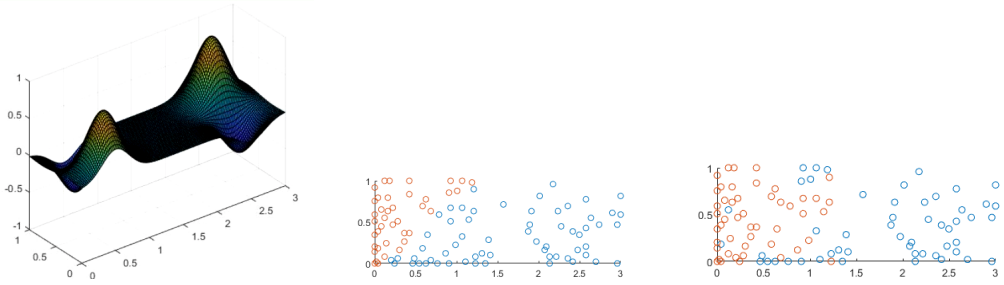
We propose the following algorithm for generating high-scoring paths:

1. Cluster the centers into invariant subspaces according to the k-invariant subspaces algorithm, with the best reasonable guess(es) for k
2. In each cluster, select a waypoint where the sum of all Koopman modes is maximized
3. Generate a path by selecting waypoints randomly, where the selection is randomly weighted. The weights are equal to the spatial extent of the clusters multiplied by an exponential decay factor for the number of times each has been visited

4 Results

4.1 Simple Synthetic Examples

To begin our investigation into this problem, we constructed a number of simple synthetic spatiotemporally evolving systems. One of these is displayed below 2(a). This is a system with two invariant subspaces, in each of which, two hills oscillate up and down in sync at a different frequency than the other invariant subspace. This is ideal for testing our methods. In the below example, we used 100 centers with an RBF kernel with a bandwidth of 0.2, however in testing we also went as low as 20 centers. The k-means algorithm divides the system nicely through the middle, as expected.



(a) Synthetic system (b) Initial clustering of centers (c) Final clustering of centers

Figure 1: Synthetic System Invariant Subspaces

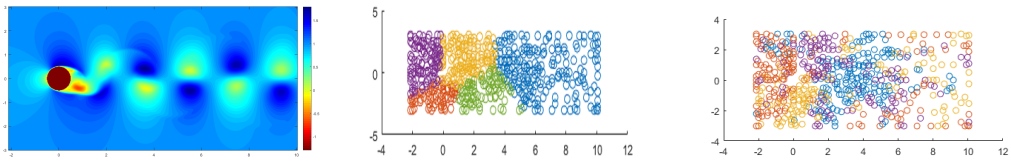
We found that few purely random paths would converge quickly for the synthetic data set, due to the large plain of zero values in-between the two invariant subspaces. Our method for generating and selecting paths with waypoints resulted in paths that traveled between the four humps and quickly converged.

4.2 Fluid Flow Past a Cylinder

We used CFD methods to generate the states for a canonical fluid mechanics problem: flow past a cylinder at various Reynolds numbers, namely Re = 100, 300, 600, 800 and 1000. This deterministic, high-dimensional spatiotemporal dynamical system is well-studied in the fluid dynamics literature, both experimentally and numerically [9, 3, 8].

For training the E-GP we used 600 centers with an RBF kernel with a bandwidth of 0.4. This output a 600×600 matrix $\widehat{A}$ which accurately captures the dynamics of the whole system.

Since the CFD set has dynamics almost everywhere, most paths are able to converge to the state. However, it is notable that a pre-designed lawnmower path failed to do as well as most random paths. Once again, the random paths generated and chosen by our methods did the best.



(a) CFD system, Re=100 (b) Initial clustering of centers (c) Final clustering of centers

Figure 2: CFD System Invariant Subspaces

5 Conclusion

In this paper we presented an a new path-planning strategy for mobile agents seeking to learn the state of a spatiotemporally evolving system, based on the spectral analysis of a linear transition matrix in a E-GP model. Our results demonstrate the viability of this approach.

References

[1] Tim P Barnett, David W Pierce, and Reiner Schnur. Detection of anthropogenic climate change in the world's oceans. *Science*, 292(5515):270–274, 2001.

[2] Joseph Berkovitz. Action at a distance in quantum mechanics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2016 edition, 2016.

[3] Chassaing P.H.H.M. Braza, M. and H.H. Minh. Numerical study and physical analysis of the pressure and velocity fields in the near wake of a circular cylinder. *Journal of fluid mechanics*, 165(130), 1986.

[4] Noel Cressie and Christopher K Wikle. *Statistics for spatio-temporal data*. John Wiley & Sons, 2011.

[5] Matthew J Heaton, Matthias Katzfuss, Shahla Ramachandar, Kathryn Pedings, Eric Gilleland, Elizabeth Mannshardt-Shamseldin, and Richard L Smith. Spatio-temporal models for large-scale indicators of extreme weather. *Environmetrics*, 22(3):294–303, 2011.

[6] H. A. Kingravi, H. Maske, and G. Chowdhary. Kernel controllers: A systems-theoretic approach for data-driven modeling and control of spatiotemporally evolving processes. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 7365–7370, Dec 2015.

[7] Hassan Kingravi, Harshal Maske, and Girish Chowdhary. Kernel observers: Systems theoretic modeling and inference of spatiotemporally varying processes. In *Advances in Neural Information Processing Systems (NIPS)*, Barcelona, Spain, 2016. accepted.

[8] Kandasamy A. Rajani, B.N. and S. Majumdar. Numerical simulation of laminar flow past a circular cylinder. *Applied Mathematical Modelling*, 33(3):1228–1247, 2009.

[9] A. Roshko. On the development of turbulent wakes from vortex streets. *California Institute of Technology*, Report 1191, 1954.

[10] Joshua Whitman and Girish Chowdhary. Learning dynamics across similar spatiotemporally-evolving physical systems. In *Conference on Robot Learning*, pages 472–481, 2017.