

Spectral learning for structured partially observable environments

Lucas Langer
lucas.langer@mail.mcgill.ca

1. PROBLEM AND MOTIVATION

We consider the problem of learning models of time series data in partially observable environments. Typical applications arise in robotics and reinforcement learning, where hidden markov models (HMMs) are often the model of choice. We take particular interest in the class of partially observable systems which are highly structured, namely systems in which one can achieve good performance with a smaller model.

There exists a well known spectral algorithm for learning a Predictive State Representation of the environment from empirical data [1]. Traditionally, truncation of learned PSRs has been performed naively. We provide an extension to [NAME]'s algorithm for improved learning of smaller models. In empirical tests, our approach strongly outperforms its ancestor in both in predictive and computational performance.

2. BACKGROUND AND RELATED WORK

Predictive state representations are used as a model for computing the probability of observations in a dynamical system. [NAME] gives an algorithm which makes use of Hankel matrices and a singular value decomposition to obtain a PSR from data. One can control the number of states in the PSR by only including states with high singular values. The reason for using less states is twofold. First, noise in empirical data artificially creates extra states with low singular values. Secondly, reducing the number of states is necessary in practice for computational performance.

Previous work with PSR learning has included planning [pierre luc] and natural language processing [other source]. Our work focuses on improving the learning of PSRs for general applications.

3. APPROACH AND UNIQUENESS

In our work, we extend the standard PSR learning algorithm by developing a new machinery for performing queries which we call the Base System. This extra machinery allows us to richly express state transitions with truncated models and provides better experimental performance. We first apply the Base System to timing applications where its construction is easiest to standardize. We then progress to the general case of systems with multiple observations and develop a heuristic for constructing the Base System effectively from data.

4. RESULTS AND CONTRIBUTIONS

In the experiments that follow, we produce observations by simulating robot motion in labyrinth environments. We compare PSRs which use the Base System to different degrees. To measure performance, the predictions of learned PSR are compared to the actual probability distribution of the observations.

4.1 Double Loops

In the first experiment we look at the time spent in double loop labyrinths.

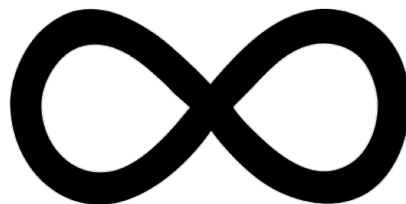


Figure 1: Double Loop Environment

In both cases, the PSR with the Base System has 100 % less error than without. In particular, we note that noise in the durations of loops doesn't harm the performance of the Base System.

4.2 PacMan Labyrinth

In the second experiment, we look at timing for a PacMan-Type labyrinth. In addition, we use state weightings from the learned PSRs to predict distances between the robot and objects in the environment.

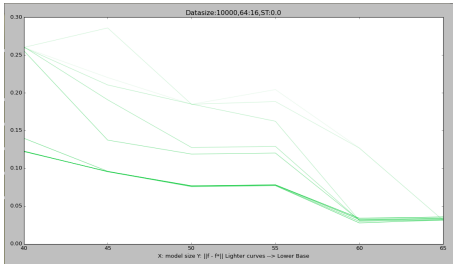


Figure 2: No noise

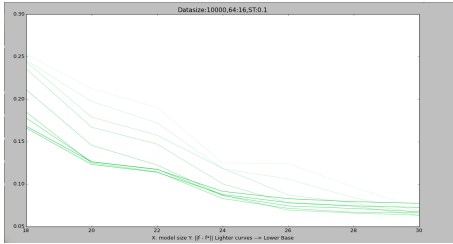


Figure 3: Corridor noise

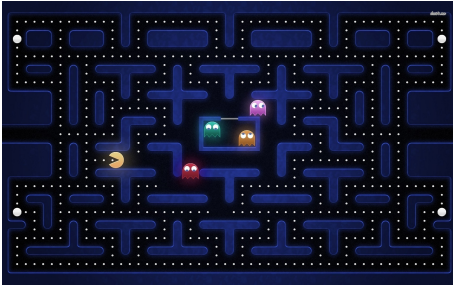


Figure 4: Pacman Environment

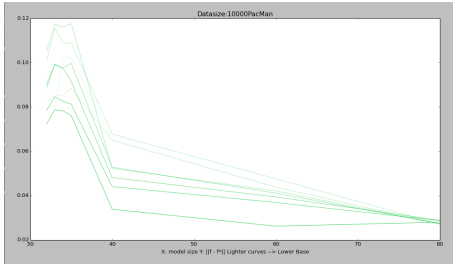


Figure 5: Timing predictions in Pacman

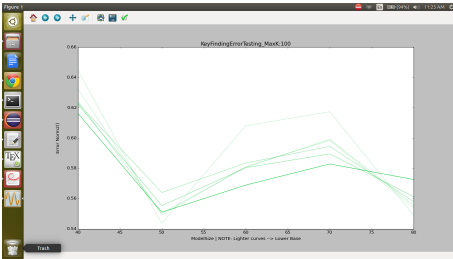


Figure 6: Distance predictions from key

Here, the Base System outperforms the naive by 100% for timing and 45% for distances.

4.3 Multiple Observations

Next, we change our set of observations to wall colors of the labyrinth.

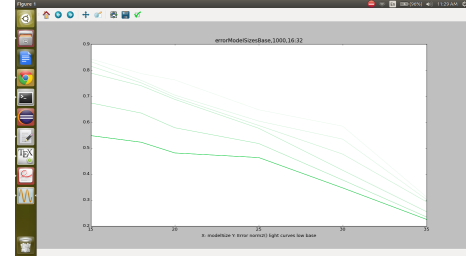


Figure 7: Predicting wall colors

Here, the Base System outperforms that naive approach by 55%. For this environment, we construct the Base System separately for each symbol. In general, one might want to use a custom heuristic to optimize the construction.

5. CONCLUSION AND FUTURE WORK

The spectral framework for learning in partially observable environments has better theoretical guarantees [REFERENCE] than non-spectral methods. In this work, we showed a way to significantly improve results in structured environments, that is when one wants a smaller model. In future work, we hope to see a theoretical analysis of the apparent improvement of the Base System and optimization of it's construction.

6. REFERENCES

- [1] M. Bowman, S. K. Debray, and L. L. Peterson. Reasoning about naming systems. *ACM Trans. Program. Lang. Syst.*, 15(5):795–825, November 1993.

6.1 References

Generated by bibtex from your .bib file. Run latex, then bibtex, then latex twice (to resolve references) to create the .bbl file. Insert that .bbl file into the .tex source file and comment out the command \thebibliography.