Bimodal Gaussian Process Transition Dynamics Models

Aidan J. Scannell
University of Bristol
aidan.scannell@bristol.ac.uk

Carl Henrik Ek
University of Bristol
carlhenrik.ek@bristol.ac.uk

Arthur Richards
University of Bristol
arthur.richards@bristol.ac.uk

***Abstract*—We present an approach for separating low and high noise regimes in Gaussian process transition dynamics models. We assume that our data is generated by two underlying processes where the assignment of a data point to a process is governed by an implicit Riemannian manifold. As such, we are able to specify a prior distribution over this *separation* manifold, providing a mechanism for encoding our expert domain knowledge regarding the separation of the two regimes in input (state-action) space. It also explicitly quantifies the uncertainty in the location of the manifold. This model was developed for application to drone operations in self induced turbulence and seeks to improve the applicability of learned dynamics models to environments that contain regimes that are subject to unmodelable stochastic effects.**

I. INTRODUCTION

In this work we are interested in dynamical systems where safety guarantees and data-efficient learning are of paramount importance. In particular, we consider dynamical systems that are subject to noise (stochastic effects) that are much higher in certain regimes of the state-action (input) space. This can be perceived as the dynamical system having two operational regimes, one that is stable and one that may be unstable (due to the presence of stochastic effects that cannot be modelled). Utilising standard control techniques with a dynamics model that has not separated the different operational regimes will have a detrimental impact on performance. We can overcome this issue by "hard coding" our domain knowledge, however, we would like our system to be able to update its belief as it receives more data through interacting with its environment (model-based reinforcement learning).

Model-based methods suffer from model bias, that is, they assume that the learned transition dynamics accurately represent the real environment. Model bias is exemplified in situations where few samples and no informative prior knowledge are available [2]. Learning transition dynamics with Gaussian processes (GP) maintains a distribution over a set of possible dynamics models, as opposed to learning a single model. Quantifying where a particular model is likely to be accurate and where regions of the underlying system are likely to be stochastic have significant consequences when considering real time control.

Although learning probabilistic dynamics models helps deal with model bias it does not address the issue of different operational regimes. Learning a dynamics model without separating the different operational regimes will have a detrimental impact on performance. We therefore seek a model that can identify and learn a separation between the different operational regimes and as a result learn a better representation of the true underlying dynamics.

In this work we seek to extend Gaussian process dynamics models to a bimodal setting, where the noise in one mode is much larger than the other. We propose to jointly learn a probabilistic separation manifold that separates the two modes in input (state-action) space along with separate Gaussian process dynamics models for each mode.

II. PROBLEM STATEMENT

In this paper we consider dynamical systems,

$$\mathbf{s}_t = f(\mathbf{s}_{t-1}, \mathbf{a}_{t-1}),$$

with unknown nonlinear transition dynamics f and continuous states $\mathbf{s} \in \mathbb{R}^D$ and actions $\mathbf{a} \in \mathbb{R}^F$. We assume that the transition dynamics f consist of two modes $\{f_1, f_2\}$,

$$f = \begin{cases} f_1 + \epsilon_1, & \text{if } h(\mathbf{s}_{t-1}, \mathbf{a}_{t-1}) < 0 \\ f_2 + \epsilon_2, & \text{if } h(\mathbf{s}_{t-1}, \mathbf{a}_{t-1}) \geq 0 \end{cases}$$

with additive, zero-mean, Gaussian system noise $\epsilon_k \sim \mathcal{N}(0, \Sigma_k)$, where the noise in one mode is much larger than the other ($\epsilon_2 \gg \epsilon_1$). $h(\mathbf{s}_{t-1}, \mathbf{a}_{t-1})$ represents an implicit Riemannian manifold that separates the two modes in input (state-action) space. Figure 1 shows a simplified dataset with a 1D input (state-action) space that represents the problem graphically. We are interested in jointly learning the separation manifold h and a dynamics model for each mode $\{f_1, f_2\}$.

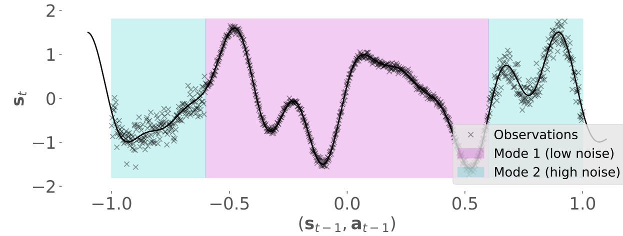

Fig. 1. Simplified dataset (1D input space) containing bimodal noise.

III. Bayesian Model

Our method models each mode using a Gaussian process dynamics model taking state and action tuples $\mathbf{x} = (\mathbf{s}_{t-1}, \mathbf{a}_{t-1}) \in \mathbb{R}^{D+F}$ as training inputs and state differences $\mathbf{y} = \mathbf{s}_t - \mathbf{s}_{t-1} + \epsilon, \epsilon \sim \mathcal{N}(0, \Sigma_\epsilon), \Sigma_\epsilon = \text{diag}([\sigma_{\epsilon_1}, ..., \sigma_{\epsilon_D}])$ as training targets [2]. We parameterise an implicit separation manifold $h(\mathbf{x})$ as a Gaussian process in order to separate the two modes. The separation manifold h assigns inputs to one of the two modes.

Our proposed model assumes that there are two independent latent functions (modes) $\{f^{(1)}, f^{(2)}\}$ that generate N pairs of observations $\{(\mathbf{x}_n, \mathbf{y}_n)\}_{n=1}^N$. The assignment of the n^{th} data point is specified by the latent variable $\alpha_n \in \{0, 1\}$. We denote all of the N inputs $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_N]$, all of the targets $\mathbf{Y} = [\mathbf{y}_1, ..., \mathbf{y}_N]$ and the assignments $\boldsymbol{\alpha} = \{\alpha_1, ..., \alpha_N\}$. The latent function value associated with the k^{th} mode and the n^{th} data point is denoted as $\mathbf{f}_n^{(k)} = f^{(k)}(\mathbf{x}_n)$ and are collected as $\mathbf{f}^{(k)} = [\mathbf{f}_1^{(k)}, ..., \mathbf{f}_N^{(k)}]$. We further denote $\mathbf{F} = [\mathbf{f}^{(1)}, \mathbf{f}^{(2)}]$. Similarly for the separation manifold h, we denote the latent function value associated with the n^{th} data point as $\mathbf{h}_n$ and collect them as $\mathbf{h} = [\mathbf{h}_1, ..., \mathbf{h}_N]$. Figure 2 shows the graphical model of the generative process. The likelihood of our model takes the form,

$$p(\mathbf{Y}|\mathbf{F}, \boldsymbol{\alpha}) = \prod_{n=1}^{N} \mathcal{N}(\mathbf{y}_n|\mathbf{f}_n^{(1)}, \epsilon^{(1)})^{\alpha_n} \mathcal{N}(\mathbf{y}_n|\mathbf{f}_n^{(2)}, \epsilon^{(2)})^{1-\alpha_n},$$

and we place independent GP priors on each of the two latent functions $\mathbf{F} = \{\mathbf{f}^{(1)}, \mathbf{f}^{(2)}\}$,

$$p(\mathbf{F}|\mathbf{X}) = \prod_{k=1}^{2} \mathcal{N}(\mathbf{f}^{(k)}|\mathbf{0}, k^{(k)}(\mathbf{X}, \mathbf{X})),$$

where we have assumed a zero mean function. We assume that there is a relationship between our assignments α_n and the input space $\mathbf{x}$ such that we can separate the assignments according to a Riemannian manifold. Placing a GP prior on this manifold means that we can encode our prior knowledge of the associations by our choice of mean function,

$$p(h|\mathbf{X}) \sim \mathcal{N}(\mathbf{h}|\mu_h(\mathbf{X}), k_h(\mathbf{X}, \mathbf{X})).$$

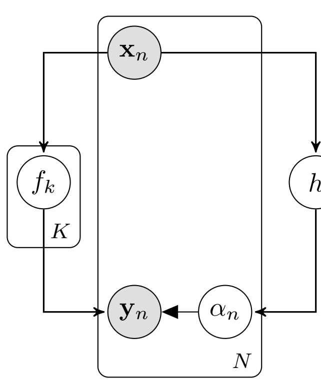

Fig. 2. Graphical model where the shaded observations $(\mathbf{x}_n, \mathbf{y}_n)$ are generated by the latent (unshaded) process.

IV. Approximate Inference

Exact inference in our model is intractable and thus we require an approximation. Stochastic variational inference (SVI) allows variational inference for very large data sets. It requires the approximation to have a set of global variables and to factorize along both the latent variables and the observations. We can compute a closed form Jensen's lower bound after expanding the joint probability model with M extra samples (auxiliary inducing points $\mathbf{u}_{j,:}$) of the GP latent mappings $\{f^{(k)}\}_{k=1}^2$ and h. As introduced in [1], we can use these extra variables in the spare GP framework from [3].

For each GP we assume that given enough well placed inducing variables $\mathbf{U}$, they are a sufficient statistic for the latent function values, implying conditional independence of $\mathbf{f}_n$ given $\mathbf{U}$ and $\mathbf{X}$. Our variational posterior takes the factorized form,

$$q(\mathbf{h}, \mathbf{F}, \mathbf{U}) = \prod_{n=1}^{N} p(\mathbf{h}_n|\mathbf{U}_h, \mathbf{x}_n) q(\mathbf{U}_h) p(\mathbf{F}_n|\mathbf{U}_f, \mathbf{x}_n) q(\mathbf{U}_f).$$

We wish to obtain a posterior over both the modes $\mathbf{F}$ and the assignments $\boldsymbol{\alpha}$. Instead of marginalising over $\boldsymbol{\alpha}$, we consider the variational joint of $\mathbf{Y}$ and $\boldsymbol{\alpha}$ to obtain a lower bound on the log joint log$p(\mathbf{Y}, \boldsymbol{\alpha}|\mathbf{X})$,

$$\mathcal{L} = \sum_{n=1}^{N} \mathbb{E}_{q(\mathbf{f}_n)q(\mathbf{h}_n)} \left[\log\, p(\mathbf{y}_n|\mathbf{f}_n, \boldsymbol{\alpha}_n) p(\boldsymbol{\alpha}_n|\mathbf{h}_n) \right]$$
$$- \text{KL}(q(\mathbf{U}_h)||p(\mathbf{U}_h|\mathbf{Z}_h)) - \sum_{k=1}^{2} \text{KL}(q(\mathbf{U}_f^{(k)})||p(\mathbf{U}_f^{(k)}|\mathbf{Z}_f^{(k)}))$$

V. Preliminary Results

Figure 3 shows our preliminary results where we have assumed that $\boldsymbol{\alpha}$ is known. It is clear that the model is capable of successfully learning appropriate variances for each mode, which is not the case when a single GP is used.

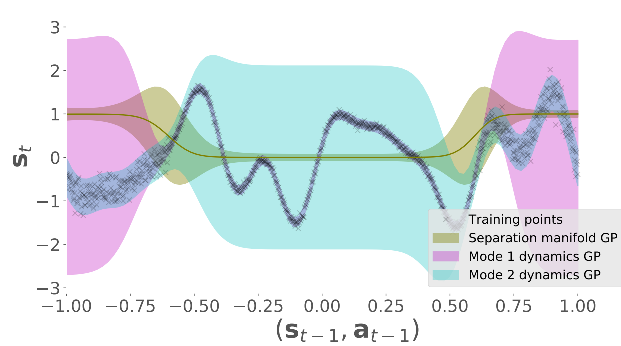

Fig. 3. Preliminary results assuming $\boldsymbol{\alpha}$ is known.

VI. Conclusion

This paper presents our initial work developing a probabilistic transition dynamics model that can accurately separate low and high noise regimes. The formulation of our model enables us to encode our prior knowledge regarding the shape of the separation manifold through the mean function of its GP prior.

Our model effectively quantifies uncertainty and enables the incorporation of prior knowledge and thus should lead to high data-efficiency. We hope that by providing this mechanism in GP dynamics models we will help advance the ability of both standard control and model-based reinforcement learning techniques to operate in environments that contain regimes that are subject to unmodelable stochastic effects.

REFERENCES

[1] Andreas Damianou. Deep Gaussian Processes and Variational Propagation of Uncertainty. Technical report, 2015.

[2] Marc P Deisenroth and Carl E Rasmussen. PILCO: A Model-Based and Data-Efficient Approach to Policy Search. *Proceedings of the International Conference on Machine Learning*, pages 465–472, 2011.

[3] Michalis K Titsias. Variational Learning of Inducing Variables in Sparse Gaussian Processes. Technical report, 2009.