Research statement

Pinak Mandal

Background

I am a final year PhD student in mathematics at International Centre for Theoretical Sciences (ICTS) - Tata Institute of Fundamental Research. Before joining PhD I obtained my BSc and MSc in mathematics from Jadavpur University. I like working on a broad range of problems, although real-world applicability is not essential, I find it to be additionally pleasing.

Some past projects: Under Prof Vishal Vasan I have worked on atmospheric plume modelling which included solving a PDE model and integrating the solution on the road network obtained from OpenStreetMap data to estimate air pollution caused by traffic around Peenya, Bengaluru. At the Indian Institute of Science I have worked on several projects during coursework which I have enjoyed immensely. Some of them are analyzing Prof Ramesh Hariharan's chromosome data and determining the most likely gene configurations that cause color blindness, finding social structures in a group of dolphins that reside in Doubtful sound, a fjord in New Zealand, music genre classification with machine learning. In 2021 I was part of Neuromatch Academy, where I worked with several other international students with varied backgrounds (ranging from neuroscience to chemistry) and we showed that fMRI data can be used to predict images seen by a subject quite well. In addition to these I have also found myself enjoying genetic algorithms and computational geometry.

Current projects: Under Prof Amit Apte's guidance I have been working non-linear filter stability and solving PDEs with machine learning. In 2022 I was selected as a part of the Future research talent program hosted by Australian National University (ANU) and have been exploring the possibility of solving magnetohydrodynamics equations with machine learning.

Nonlinear filter stability

In earth sciences and various other fields a great amount of data is constantly being collected and an important problem is to incorporate these data into physical dynamical models for making real time predictions for weather forecasts, hurricane tracking to name a few. This problem known as data assimilation is often tackled with a Bayesian approach which naturally leads to the problem of nonlinear filtering [1, 2], which studies the conditional distribution, called the filter or the posterior distribution, of the state at any time conditioned on observations up to that time [12, 7].

In case of deterministic dynamics the problem of non-linear filtering can be stated as follows. We have a state-vector $x_k \in \mathbb{R}^d$, at discrete time k, modelled to follow

$$x_{k+1} = f(x_k) \tag{1}$$

which we have access only through, often lower dimensional, noisy observation

$$y_k = Hx_k + \eta_k \tag{2}$$

where H is a matrix of dimension $q \times d$ and $\eta_k \sim \mathcal{N}(0_q, \sigma^2 I_q)$ are i.i.d. additive Gaussian noises. We also have a guess for the initial state x_0 as a distribution $x_0 \sim \mu$. The goal of filtering is to determine $\pi_n(\mu) := p(x_n|y_{0:n})$ or the distribution of x_n given all the observations up to time n. In practice, our guess for the initial distribution μ might be far off from the actual initial state vector. A measure of robustness of a filtering algorithm is how well it is able to "forget" the initial distribution i.e. given two different initial distributions does a filtering algorithm produce near-identical filtering distributions after a while?

We propose a new definition to assess this property called "stability" of a filter and propose a computationally tractable way to compute it. Using this definition we demonstrate two popular Monte-Carlo based filtering algorithms the particle filter [5] and the ensemble Kalman filter [6] are stable in chaotic test cases [8]. In a separate work [9] I show that under certain conditions the new definition is a stronger version of previous stability definitions [11] which despite having theoretical appeal lack methods of direct computation. We also establish numerically a relation between filter stability and filter convergence by showing that the Wasserstein distance between filters with two different initial conditions stays proportional to the bias or the RMSE of the filter before reaching stability.

Future work

The linear relationship between filter RMSE and filter stability can be seen experimentally but a mechanism which causes this remains unclear and hence is worth exploring.

Solving Fokker-Planck equations with machine learning

Fokker-Planck equations describe evolution of distributions under noisy dynamics and present themselves in a wide variety of topics from granular media [3] to plasma physics [10]. In this (finished but unpublished) work we focus on Fokker-Planck equations of the form,

$$\frac{\partial p}{\partial t} = \mathcal{L}p \stackrel{\text{def}}{=} -\sum_{i=1}^{d} \frac{\partial}{\partial x_i} (\mu_i p) + \sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2}{\partial x_i \, \partial x_j} (D_{ij} p) \tag{3}$$

$$= -\nabla \cdot (\mu p) + D \odot \nabla^2 p \tag{4}$$

where $\mu : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}^d$ is a vector-valued and $\sigma : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \to \mathbb{R}^{dl}$ is a $d \times l$ matrix-valued function Here $\odot$ is the Hadamard product, ∇^2 implies Hessian and $D = \frac{1}{2}\sigma\sigma^\top$. For ease of discussion we classify systems into two broad categories gradient systems where $\mu = -\nabla V$ for some potential function V and non-gradient systems where the drift μ can not be expressed in such a way.

I establish a physics informed neural net based method to solve for unnormalized steady-states in which just by carefully avoiding one local minimum of the loss function it is possible to compute unnormalized steady states even for dimensions traditionally condsidered to be challenging eg $d = 10$. I point out that even though our networks are dense in a Sobolev space containing the true solutions, a similar method for the time-dependent equation is not guaranteed to converge to the true solution in practice. To deal with this I solve an auxiliary backward Kolmogorov equation with Feynman-Kac formula and combining it with unnormzalized steady-state I compute the normalized time-dependent solution. I focus on an important class of systems (both gradient and non-gradient) that possess an attractor and explain that the final method is capable of computing time-dependent solutions for gradient systems upto an arbitrary amount of time and for non-gradient systems upto a finite amount of time.

Future Work

The Bayesian filtering problem is usually solved through two steps. In the first step called "prediction" we compute the time evolution of the distribution at the previous time-step. In case of noisy dynamics it is akin to solving a Fokker-Planck equation and therefore a new filtering method can be imagined that uses a Fokker-Planck solver. Whether the method described here can be extended to an arbitrary amount of time for non-gradient systems is another topic of interest.

Solving variational problems with machine learning

In this ongoing work I demonstrate that variational problems can be solved with machine learning and as an example I compute the minimal surface given a boundary in $\mathbb{R}^3$. I also explore three different ways of satisfying constraints in a machine learning setting - by design, by the penalty method and by the augmented Lagrangian method [13] and experimentally show that the augmented Lagrangian method takes fewer iterations of gradient descent than the penalty method to converge at the cost of using more memory.

Future work

I, in collaboration with the plasma physics group at Australian National University led by Prof Matthew Hole, am currently exploring the possibility of solving magnetohydrodynamics equations written in variational form [4] using the augmented Lagrangian method and machine learning.

A gallery of examples

A gallery of examples from my current projects is available at my Github page https://pinakm9.github.io/gallery22/.

References

[1] Mark Asch, Marc Bocquet, and Maëlle Nodet. *Data assimilation: methods, algorithms, and applications.* SIAM, 2016.

[2] Alberto Carrassi, Marc Bocquet, Laurent Bertino, and Geir Evensen. Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, 9(5):e535, 2018.

[3] José A Carrillo, Robert J McCann, and Cédric Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Revista Matematica Iberoamericana*, 19(3):971–1018, 2003.

[4] RL Dewar and ZS Qu. Relaxed magnetohydrodynamics with ideal ohm's law constraint. *Journal of Plasma Physics*, 88(1), 2022.

[5] Arnaud Doucet, Adam M Johansen, et al. A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of nonlinear filtering*, 12(656-704):3, 2009.

[6] Geir Evensen et al. *Data assimilation: the ensemble Kalman filter*, volume 2. Springer, 2009.

[7] Kody Law, Andrew Stuart, and Konstantinos Zygalakis. Data assimilation. *Cham, Switzerland: Springer*, 214:52, 2015.

[8] Pinak Mandal, Shashank Kumar Roy, and Amit Apte. Stability of nonlinear filters-numerical explorations of particle and ensemble kalman filters. In *2021 Seventh Indian Control Conference (ICC)*, pages 307–312. IEEE, 2021.

[9] Pinak Mandal, Shashank Kumar Roy, and Amit Apte. Probing robustness of nonlinear filter stability numerically using sinkhorn divergence. *arXiv preprint arXiv:2208.10810*, 2022.

[10] Arthur G Peeters and Dafni Strintzi. The fokker-planck equation, and its application in plasma physics. *Annalen der Physik*, 17(2-3):142–157, 2008.

[11] Anugu Sumith Reddy and Amit Apte. Stability of non-linear filter for deterministic dynamics. *arXiv preprint arXiv:1910.14348*, 2019.

[12] Sebastian Reich and Colin Cotter. *Probabilistic forecasting and Bayesian data assimilation.* Cambridge University Press, 2015.

[13] Stephen Wright, Jorge Nocedal, et al. Numerical optimization. *Springer Science*, 35(67-68):7, 1999.