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. So far I have enjoyed working on a broad range of applied mathematics or computational problems. Some of them are given below.

- Under prof Vishal Vasan I have worked on air pollution modelling which used a simple PDE model for atmospheric plume and incorporating real world data into the model to compute pollutant concentrations.
- At the Indian Institute of Science I have worked on several projects. A few 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 dolphins of Doubtful sound, a fjord in New Zealand,
- music genre classification with machine learning,
- finding approximate solutions to computationally challenging with genetic algorithms.
- In 2021 I was part of Neuromatch Academy, where I worked with several other international students with varied backgrounds (ranging from neuroscience to chemistry) to show that fMRI data can be used to predict types of images seen by human subjects.
- I have written libraries for simulating general purpose simulation of stochastic processes, Markov chains with implementations of popular filters often used in data assimilation.
- Under Prof Amit Apte's guidance I have been working on data assimilation (focusing on numerical filter stability) and solving high dimensional PDEs (tested in 10 dimensions).
- 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 written in an augmented Lagrangian form under prof Matthew Hole.

Nonlinear filter stability

In earth sciences and various other fields a great amount of data are 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 is known as data assimilation (DA). Bayesian approach to DA leads to the problem of nonlinear filtering [1, 2], which studies the conditional 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 the distribution of x_n given $y_{0:n}$, all the observations up to time n which is called the filtering distribution. 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 using the Wasserstein metric. 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.
- Since in the presence of attractors the filtering distributions are eventually supported on the attractor, sampling the attractor might lead to more robust filtering algorithms. Also, gradient descent in Wasserstein distance might facilitate better sampling of the posterior which is another point of interest.

Solving high dimensional Fokker-Planck equations

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 work (in preparation) we focus on Fokker-Planck equations of the form,

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

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

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 $\Sigma = \frac{1}{2}\sigma\sigma^\top$. 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 PINN (physics informed neural net)-based algorithm that carefully avoids one critical point of the loss function and therefore is able to compute unnormalized steady states even for dimensions traditionally condsidered to be challenging eg $d = 10$. Even though the networks are dense in Sobolev spaces containing the true solutions, similar methods for the time-dependent equation are not guaranteed to converge to the true solution in practice. To deal with this an auxiliary backward Kolmogorov equation can be solved with Feynman-Kac formula and combining it with unnormzalized steady-state one can 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 a topic of interest.

- The non-convex structure of the loss function is another point of interest. All the critical points of the loss function for the steady state might lie on a connected one dimensional manifold or might be separated from each other. And this sturcture can be probed with numerical experiments.

Solving constrained variational problems with machine learning

Constrained variational problems are ubiquitous in physics due the stationary-action principle. Here I look at problems of the following form,

$$\underset{u(\mathbf{x})}{\text{minimize}} \int_{\Omega} J(\mathbf{u}, D\mathbf{u})\, d\mathbf{x} \quad \text{subject to} \quad C(\mathbf{u}(\mathbf{x}), D\mathbf{u}(\mathbf{x})) = 0 \;\forall\, \mathbf{x} \in \Omega \quad \text{and} \quad \mathbf{u}|_{\partial\Omega} = g \tag{5}$$

where D denotes derivatives. Finite dimensional constrained optimization problems can be tackled with the augmented Lagrangian method [13]. In this method one iteratively updates one's guess for both the solution and the Lagrange multipliers. Since in (5) we are looking for a function or attempting to solve an infinite dimension optimization problem, machine learning can be employed as a useful tool. Just like the finite dimensional optimization problems, we can compute both the solution and the Lagrange multipliers represented as neural nets with an alternate gradient descent scheme.

In this ongoing work I demonstrate as an example that with this method, minimal surfaces like helicoid (non self-intersecting) and Enneper's surface (self-intersecting) can be computed given the boundary data in $\mathbb{R}^3$. This problem can be set up as,

$$\underset{u}{\text{minimize}} \iint_{\Omega} \sqrt{1 + u_x^2 + u_y^2}\, dx\, dy \quad \text{subject to} \quad u|_{\partial\Omega} = g \tag{6}$$

In this case the associated Euler-Lagrange PDE given by,

$$(1 + u_x^2)u_{yy} - 2u_x u_y u_{xy} + (1 + u_y^2)u_{xx} = 0 \;\; \forall\, (x, y) \in \Omega \quad \text{subject to} \quad u|_{\partial\Omega} = g \tag{7}$$

can be also be solved using machine learning with the added cost of the extra derivative computation. 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. I experimentally show that the augmented Lagrangian method converges much faster than the penalty method at the cost of using more memory. For this experiment I use a toy problem of computing the magnetic field in a 3D box, given the boundary data by satisfying Gauss's law, Ampere's law and minimizing energy.

Future work

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