Research Statement

Shashank Kumar Roy
PhD Research Scholar, International Center for Theoretical Sciences, Bangalore

January 2023

Overview

I have pursued my research work in the area of data assimilation before which I finished my M.Sc and B.Sc Honors in physics. During my coursework, I had the chance to work on interesting projects starting with analytical solution of Stokes's flow in spherical geometry using vector spherical harmonics, numerical solution to stochastic forced burgers equation, 1-d visco-eslastic pde model, kuramoto-sivashinsky equation and ensemble kalman filtering for Lorenz-96, a 40-dimensional chaotic dynamical system using partial and noisy observations.

I am generally fond of problems in prediction and inference in context of complex systems where well defined mathematical notions and ideas of probability and statistics provide insight and can be converted to code and increase my understanding through simulations. Other endeavours orthogonal to my research work which piqued my curiosity from time to time were participation in hackathons where I often stumble across new problems and fascinating progress in deeplearning methods and their applications. I have taken a few courses in machine learning and deeplearning from neuromatch academy and nvidia workshops. Most of my coding experience in machine learning with real datasets was gained during hackathons organized from scientific communities such as NOAA and Ecole Polytechnique, such as writing an LSTM model for prediction of spatial time series data of weather variable and deep generative models for learning distribution of sea surface temperature from historical data using a generative adversarial network(GAN).

My thesis research work concentrates around data assimilation for chaotic dynamical system using EnKF[1], a general sequential state estimation algorithm which computes the best estimate of the state with associated uncertainty. The underlying theme of my interests have been in studying filtering algorithms and their properties which can be used to diagnose and improve the same. Jointly with others, I have worked on demonstrating numerical filter stability, a crucial property of a filter using Sinkhorn distance, a distances between probaility distribution. In another work, I am looking at instability properties of a dynamical system such as lyapunov vectors which are important in improving the existing techniques in prediction and estimation of a dynamical system in general. Now I briefly talk about them below, starting with filter stability.

Numerical filter stability

In high-dimensional chaotic system such as atmosphere and ocean, it is impossible to track and predict such a system for long using only the model since amplification of various sources of errors such as error in initial condition, parameters and model error lead to rapid departure of the two nearby solutions [2]. Data assimilation addresses two problems of inverse modeling which are filtering and prediction, where numerical models of the physical system representing our knowledge of the governing dynamics are combined with the noisy and sparse observations from the system weighted by their respective uncertainties, in order to produces improved statistical estimates of the true state of the system[3]. The best estimates are then used to predict future observation. In the setting of a deterministic dynamical system, with observations operators h and measurement noise ϵ, we have

$$\mathbf{x}_k = \mathbf{M}(\mathbf{x}_{k-1}), \mathbf{y}_k = h(\mathbf{x}_k) + \epsilon_k, \epsilon_k \sim \mathcal{N}(0, \Sigma) \tag{1}$$

where, $\mathbf{x}_k \in R^d, \mathbf{y}_k \in R^p$ and $p < d$.

Bayesian filtering algorithms for data assimilation

Bayesian filtering is defined as the sequential estimation of the conditional distribution in phase space of the state of a physical system coming from an assumed model taking into account the likelihood of new information arriving from the observations [4]. Filtering is followed by prediction, where the goal is to predict the the future of the system and it's flow dependent uncertainty accounting for possible source of error to a later time in future. If At time k , the probability density distribution of the true state conditioned on observations upto time k is called the analysis probability density, given by $\rho\left(\mathbf{x}_k|\mathcal{Y}_k\right)$. This is then propagated in time by solving fokker-planck equation, or by using monte-carlo methods [1]to integate the ensembles representing the distribution to obtain prior distribution at time $k+1$, called forecast probability density distribution at time $k+1$, before using the information about the observation $\mathbf{y}_{k+1}$. Assuming that given x_{k+1}, the observations y_{K+1} is conditionally independent of all previous observations and states, we get the likelihood for the observation $\mathbf{y_{k+1}}$ is obtained using the observation model and the measurement noise distribtution is given by $\rho\left(\mathbf{y}_{k+1}|x_{k+1}\right)$. A bayesian filtering algorithm begins with an initial distribution $\rho(x_0)$ of the state at $k=0$, which could be far from the true unknown initial state and for each time k when the observation arrives, using bayes theorem, one can write,

$$\rho\left(\mathbf{x}_{k+1}|\mathcal{Y}_{k+1}\right) = \frac{\rho\left(\mathbf{y}_{k+1}|\mathbf{x}_k+1\right)\rho\left(\mathbf{x}_{k+1}|\mathcal{Y}_k\right)}{\rho\left(\mathbf{y}_{k+1}|\mathcal{Y}_k\right)} \tag{2}$$

Stability of a filtering algorithm

A numerical filter such as EnkF and particle filter, approximates numerically the conditional distribution of the state starting with observation $y_{0:K}$ and a initial distribution of the true state at $k=0$, sequentially over time using ensemble approach. Different filtering algorithms are based upon different approximations making the bayesian posterior distribution computations tractable and using the flow-dependent uncertainty provided by the governing dynamical equations of the system along with a model of observations schemes of updating with the uncertainty of the model estimates using the observation y_k [4].

Since the true state is unknown, all filtering algorithms makes an arbitrary choice for $\rho(\mathbf{x_0})$ at the time of filter initialization. It becomes crucial that these distributions become independent of the initial choice used in initializing the filter so that quantities estimates using the posterior are eventually independent of our arbitrary and often wrong choice of $\rho(\mathbf{x_0})$. Filter stability is the property that the posterior distribution computed sequentially over long time is independent of the choice of the distribution at $k=0$ used to initialize the filtering algorithm. The rate at which the filter achieves this independence, the better are the estimates of various quantities which utilize the posterior distribution. Mathematically, we can define it using a distance D on the space of probability distributions $P(\mathcal{R}^d)$.

Given two different initial distributions ν_1, ν_2 for x_0 with $\hat{\pi}_n(\nu_1)$ and $\hat{\pi}_n(\nu_2)$ being the posterior obtained after assimilating all observation $y_{1:n}$ at time n, then for filter stability to hold, we have

$$\lim_{n\to\infty}\mathbf{E}[D(\hat{\pi}_n(\nu_1),\hat{\pi}_n(\nu_2))] = 0 \tag{3}$$

where, D is a distance on $P(\mathcal{R}^d)$, the space of probability measures on $\mathcal{R}^d$ and the expectation is over observation noise accounting for different noise realizations [5]. In our work, we have chosen Sinkhorn distance, a distance well motivated by the ideas of optimal transport, which has several merits over other distances such as total variation, apart from being computationally cheaper. This distance utilizes two samples from their respective distributions in order to compute the distance between them upto it's statistical errors. For different test cases, I then show that EnkF is stable by empirically computing the the distance between the ensembles representing the conditional distribution over time. It also allows us to study rate of the filter stability.

Future research questions

Below are some of problems where application of the above ideas of filter stability may be practically useful.

1 A particularly interesting idea is to see how probabilistic machine learning models proposedly performing filtering, satisfy this criteria improving them.

2 The study can be generalized to different filters in context of model and parameterization errors of unmodelled quantities respectively.

3 Relating filter stability to the chaotic properties of the underlying system, since the interplay between the instability and the informative observations lead to eventually capture the conditional distribution.

Computation of Covariant Lyapunov vectors

Instability properties of a chaotic nonlinear dynamical system are characterized globally by lyapunov exponent and locally by lyapunov vectors, which correspond to the directions associated with those rates. Unstable subspaces of a system has been shown to improve forecasts in ocean-atmosphere coupled models when the assimilation takes into account [6]. When a time evolving system is only accessible via noisy and sparse observations and with no information of the initial conditions, identifying a trajectory is extremely difficult using only the model and methods of nonlinear filtering provide best estimates of the true state and help us estimate the underlying trajectory the system evolves on with some uncertainty estimate. The best estimates over time does not constitute a dynamical trajectory and but is always near the true trajectory in some distance metric. We formulated this problem into using the filter estimate over time as a proxy of the true trajectory perturbed by the error statistics.

Sensitivity of Lyapunov vectors

The specific questions which we ask about lyapunov vectors and exponents in context of the filtering problem are as follows:

- How sensitive are the backward and covariant lyapunov vectors and the corresponding exponents to perturbations in the underlying trajectory?
- Under what conditions can one recover them from a filter estimated trajectory instead of the true trajectory of the dynamical system?
- How robust are unstable subspaces to the perturbation strength σ, and are they more robust than the individual vectors themselves?

The above questions are relevant to both nonlinear dynamics and weather forecasting techniques since the underlying true trajectory is not accessible but are required in order to compute the LVs. What one can do best is to use the best estimate of the true state given observations upto a certain time. This comes with a caveat that the filter estimates or the analysis mean over time is not a dynamical trajectory of the system, i.e. there is not initial condition on the model which can generate this trajectory starting from a certain initial condition. The true trajectory lies near the analysis mean which is quantifies by the error in the estimates themselves. The sensitivity of different lyapunov vectors such as backward and covariant lyapunov vectors for a particular system tells us the limitations of what can be reconstructed from the filter estimates themselves.

In order to start answering the above questions, I studied how sensitive are the lyapunov vectors to the perturbations added to the underlying trajectory, which allows us to understand the limitations of such vectors obtained from the numerical state estimates of the filter. I studied this sensitivity by systematically adding noise of strength σ following a gaussian distribution $\mathcal{N}(0, \sigma^2 I_d)$, where d is the dimension of the state, to underlying true trajectory and studying their effect on the computed backward and covariant lyapunov vectors using Ginelli's algorithm[7]. I used Lorenz-63, where visualization is straight forward to interpret 1.

Covariant lyapunov vectors have been shown to predict regime change in L63 system [8], and 1 shows that through the cosines of the angle between the first two clvs. In high-dimensions, I studied Lorenz-96 in order to explore dimensional dependence added to the sensitivity problem. Another interesting directions is using principle angles([9]), which summarize the angle between two different sub-spaces, and seeing how they change with σ, which I found to be more robust than the individual vectors themselves. Such analysis is useful in context of problems where sub-spaces are more important than the individual vectors themselves. It also allows one to understand and interpret how much information can be reconstructed using lyapunov vectors from a filter generated trajectory using partial and noisy observations and a model.

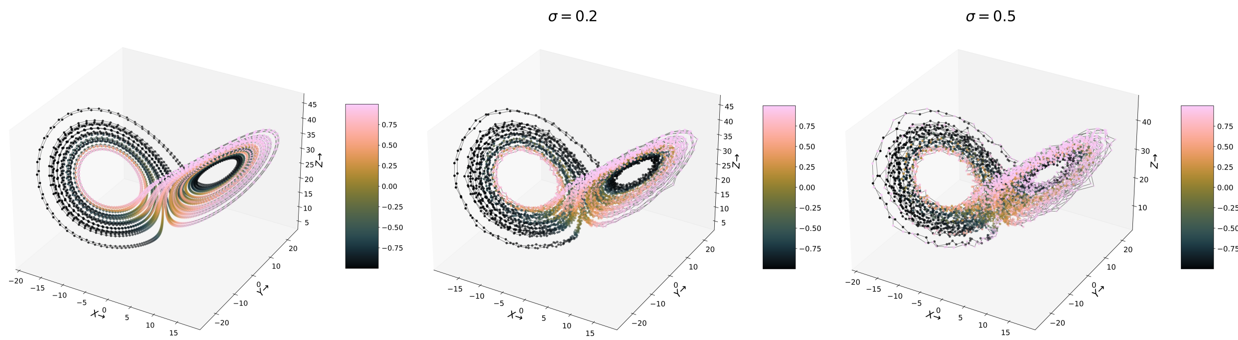

Figure 1: Plotting the cosine of the angle between the first two clvs, for both true trajectory and perturbed trajectory in phase space. The left-most picture corresponds to the true trajectory.

Possible directions for future work

A set of directions for future work directly following the current work are:

1 Combining the ideas of assimilation in unstable subspace where the sub-spaces computed from the historical data can be employed.

2 Studying the effect of model errors themselves on the computed lyapunov vectors.

3 Using the sub-spaces computed from reanalysis data which can be used to restrict the uncertainty in forecasts and analysis in data assimilation cycles.

4 The degree of similarity in the lyapunov vectors between two nearby points on the attractor in phase space which may be important machine learning methods trained on vectors computed once along a trajectory be used to predict vectors at neaby points in the phase space.

5 Studying structure of CLVs for discretized PDE systems such as Kuramoto-sivashinksy equation which has a finite dimensional attractor to shed light on CLV localization problem.

Other research interests

Problems at the interface of climate and data science is something I am interested to work, the topic on which my interest from different workshops and discussions where I believe that the skills which I have acquired in the context of data assimilation are useful. I am interested in understanding ways to incorporate uncertainty and dynamical knowledge together to a general machine learning techniques for modelling and inference of large dynamical systems in real applications such as climate and weather models where the limited data combined with physical constraints and conservation laws can balance for the sparsity and scarcity of available data. Such problems are of high interest in climate modeling and related data science problems. Another important topic which I find fascinating to explore further is optimal transport which I gained some exposure to while working on the filter stability problem and would like implement it for different contexts. Understanding generative modelling to design more complex filters, which are focused on real applications is another of my interests.

References

[1] G. Evensen, "The ensemble kalman filter: theoretical formulation and practical implementation," *Ocean Dynamics*, 2003. [Online]. Available: https://doi.org/10.1007/s10236-003-0036-9

[2] T. N. Palmer and L. Zanna, "Singular vectors, predictability and ensemble forecasting for weather and climate," *Journal of Physics A: Mathematical and Theoretical*, vol. 46, no. 25, p. 254018, jun 2013. [Online]. Available: https://doi.org/10.1088/1751-8113/46/25/254018

[3] B. M. B. L. E. G. Carrassi, A., "Data assimilation in the geosciences: An overview of methods, issues, and perspectives," *WIREs Clim Change*, vol. 75, no. 1B, pp. 257–288, 2018. [Online]. Available: https://doi.org/10.1002/wcc.535

[4] S. E. Cohen, "An introduction to estimation theory," *Journal of the Metereological Society of Japan*, vol. 75, no. 1B, pp. 257–288, 1997. [Online]. Available: https://doi.org/10.2151/jmsj1965.75.1B$_2$57

[5] P. Mandal, S. K. Roy, and A. Apte, "Stability of nonlinear filters-numerical explorations of particle and ensemble kalman filters," in *2021 Seventh Indian Control Conference (ICC)*. IEEE, 2021, pp. 307–312.

[6] S. Vannitsem and V. Lucarini, "Statistical and dynamical properties of covariant lyapunov vectors in a coupled atmosphere-ocean model-multiscale effects, geometric degeneracy, and error dynamics," *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 22, p. 224001, may 2016. [Online]. Available: https://doi.org/10.1088/1751-8113/49/22/224001

[7] F. Ginelli, H. Chaté, R. Livi, and A. Politi, "Covariant lyapunov vectors," *Journal of Physics A: Mathematical and Theoretical*, vol. 46, no. 25, p. 254005, jun 2013. [Online]. Available: https://doi.org/10.1088/1751-8113/46/25/254005

[8] E. L. Brugnago, J. A. C. Gallas, and M. W. Beims, "Predicting regime changes and durations in lorenz's atmospheric convection model," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 30, no. 10, p. 103109, 2020. [Online]. Available: https://doi.org/10.1063/5.0013253

[9] G. H. Golub and H. Z. ha, "The canonical correlations of matrix pairs and their numerical computation," 1992.