Computation of covariant lyapunov vectors using data assimilation

Shashank Kumar Roy[1] and Amit Apte[2]

[1]International Centre for Theoretical Sciences, Bangalore 560064 India

[2]Indian Institute of Sciences, Education and Research, Pune 411008 India

Correspondence: Shashank Kumar Roy (shashank.roy@icts.res.in)

Abstract. We study the sensitivity of backward and covariant lyapunov vectors to the perturbations added to the underlying orbit of the dynamical system Lorenz-96 for different perturbation strengths. Since the true trajectory which is unknown, we relate the nature of perturbed trajectory to analysis mean trajectory obtained from a filter with partial and noisy observations and study how close are the lyapunov vectors from the analysis mean trajectory to the actual vectors. We also study the principal angles between the subspace spanned by the vectors computed from the true and the analysis mean trajectory. This directly provides us a way to practically understand the limitations of the reanalysis data and their utility in context of computing these vectors for improving operational data assimilation and forecasting.

1 Introduction

Trevisan-Uboldi 2004 JAS: AUS and Observation–Analysis–Forecast Cycle https://journals.ametsoc.org/view/journals/atsc/61/1/1520-0469_2004_061_0103_aosato_2.0.co_2.xml

Lyapunov exponents and vectors are the the fundamental tools of lyapunov analysis in the context of nonlinear and chaotic system, where they describe the instability properties of a nonlinear dynamical system with respect to perturbations along different directions. For a finite n-dimensional dynamical systems where ergodic hypothesis holds, there are n lyapunov exponents which globally summarize the asymptotic growth rate of perturbations over the whole attractor, whereas lyapunov vectors are local objects which span the tangent manifold for a specific point in the phase space and contains the information of it's past and future evolution (Eckmann and Ruelle (1985)). The lyapunov exponents have been extensively computed using Benettin's algorithm for various dynamical systems(Benettin (1980)) (https://link.springer.com/article/10.1007/BF02128237#article-info), where a set of initially orthogonal vectors are integrated forward in time via the tangent linear dynamics to obtain their growth rates over small time intervals after which they are periodically re-orthonormalized using gram-schmidt orthonormalization as a necessary step to avoid numerical overflow. Numerically, the lyapunov exponents are obtained as a result of averaging the growth-rates of the perturbations over small time intervals along the trajectory. These perturbation vectors after undergoing repeated forward integration and gram-schmidt re-orthonormalization are called forward gram-schmidt vectors which, after a long time converge to what are known as the Backward Lyapunov Vectors or BLVs. BLVs signify directions in the tangent space at a point along which errors show maximal error growth under time reversed-dynamics(proper def). The forward and backward lyapunov vectors are orthogonal by construction but are not covariant under the tangent linear dynamics.

Covariant lyapunov vectors, derive their name from their most important property of being covariant under the tangent linear equations. They are norm-independent and time invariant local basis vectors in the tangent space along which perturbations grows or decays along the trajectory with the associated rates being the lyapunov exponents i.e. an infinitesimal error along i^{th} CLV grows with $exp(\lambda_i t)$ forward in time and decays as $exp(-\lambda_i t)$ backwards in time, λ being the corresponding exponent(Pazó et al. (2008),Yang et al. (2009)). Unlike the forward and backward lyapunov vectors which are orthogonal by construction, CLVs do not form a orthogonal set of vectors, which means that two CLVs can have significant projection along each other. Two major algorithms for their computation, one developed by Wolfe and Samelson (Wolfe and Samelson (2007)), and a more recent one called dynamic algorithm of Ginelli (Ginelli et al. (2013)) have been used to compute CLVs for PDEs models such as Rayleigh Benard Convection, Coupled Map Lattice, Kuramoto Sivashinksy and Lorenz-96 system. In the present work, we use Ginelli's algorithm which has been used extensively and it's convergence results are available in Noethen (2019).

Numerical weather prediction problems face the problems related to highly nonlinear and chaotic systems with huge number of degree of freedoms. Predictability issue of high-dimensional chaotic systems such as ocean and atmospheric models are limited due to large number of different sources of errors in dynamics, errors in initial conditions which grow exponentially due to their inherent chaotic nature. It is not possible to predict the state of the system beyond a certain time due to highly sensitive to the errors in the initial conditions, model errors and errors in parameters and initial conditions combined result in departure of the model state from being a close representation of the true state. To keep the state present on the numerical model close to the reality over time, one uses data assimilation where real observations which are noisy and sparse both in space and time, are used to correct the state estimated by the model to bring it close to the true state under some metric. Data assimilation methods then carry out prediction-correction step sequentially over time which generates states which are near the true state of the system. However, they do not constitute a dynamical trajectory as there is no initial condition at t=0 time, which when integrated forward in time passes though the estimated states generated sequentially.

For a dynamical system evolving in real time where the full state is not observed or is observed partially and indirectly, bayesian formalism combines the model outputs with the serially arriving observations allowing one to reasonably estimate the state at time t along with their related uncertainty. Such state estimation algorithms are called filters in the filtering theory literature and their applications is popularly known as data assimilation in the earth sciences literature. Filtering can incorporate both the cases of noisy and partial observations, which are a big challenge compared to only data driven approaches. Two popular class of such filters are ensemble kalman filters and particle filters.

Being a non-orthogonal basis, CLVs have quite distinct properties from the other type of LVs and their alignment have shown to predict the regime change in Lorenz-63, a three dimensional chaotic model Brugnago et al. (2020) https://aip.scitation.org/doi/10.1063/5.0013253. Their application in identifying ensemble methods in data assimilation is promising as optimal ensembles used in forecasting should span the space of error growth.(cite). However, we take the approach of using data assimilation to compute the CLVs using model and the observations coming from the system in a non-intrusive manner.

The uncertainty of their future predictions depend on the directions of the error growth and this information is crucial to increase the predictability as the dimension of the unstable growth is much smaller compared to the full state space which holds for high-dimensional chaotic and dissipative systems. The work of Trevisan and Palatella introduced AUS (Trevisan et al. (2010))(https://sci-hub.se/10.1002/qj.571) and Palatella et al. (2013) https://iopscience.iop.org/article/10.1088/1751-8113/46/25/254020, a paradigm where the unstable subspace spanned by the positive and neutral lyaunov vectors is used for producing the analysis or the update step of assimilation and shows promising improvements over the traditional algorithms. Bocquet and Carrasi,Bocquet et al. (2017) followed by Gurumoorthi, Apte and co-workers Gurumoorthy et al. (2017) https://doi.org/10.1137/16M1068712 have shown analytically for Kalman filter and Kalman Smoother and numerically for their ensemble version that their asympotitc covariance matrices collapse on the upstable subspace of the dynamics. The application of the LVs have recently been applied by Vannitsem and Lucarini (2016)https://link.springer.com/article/10.1007/s00382-020-05313-3 to data assimilation for ocean-atmosphere model where the initial ensemble was projected along a set of BLVs and it was seen that both unstable and slightly negative lyapunov vectors are important, insead of only the unstable ones for multi-scale coupled systems.

Papers need to be cited about predictablity- Tim Palmers,

2 Problem Statement

Grounded in the principles of filtering theory, data assimilation is a well-know technique in geosciences, which combines partial and noisy observations from a system with it's numerical model in an optimal way to estimate density distributions of the full state of the system Carrassi et al. (2018). Since estimating the full density is impossible in high-dimensions, important statistics describing the pdf are estimated such as mean and covariance and how they evolve in time. Various data assimilation algorithms differ in their specific assumptions and approximations in order to optimally combine the model and the observations making the sequential estimation of such statistics computationally tractable in high-dimensions.

Calculation of LVs require two key numerical procedures, a choice for the fixed underlying orbit of the dynamical system and second is evolving a set of perturbations in tangent space done by a the model and it's jacobian. Methods to compute these vectors numerically require a reference trajectory of the dynamical system about which the above operations are carried out. The numerical solution generated from a numerical model, as discussed earlier, departs from the true trajectory very fast due to chaotic nature of the system as the initial condition can never be known precisely. The true state may only be observed indirectly though measurements of other state-dependent physical quantities. Data assimilation ameliorates this problem by producing estimates of the state over time which are close to the true trajectory in some desired metric. These vectors and their subspaces, when computed in practice from the estimated trajectories, which are obtained from a data assimilation. Any algorithm, filtering or smoothing, which determines a estimated trajectory is not a dynamical trajectory of the model itself.

The set of best estimates over time do not constitute a dynamical trajectory i.e. there is not initial condition at t=0 such that if integrated forward in time contains the analysis means obtained over time. Yet it is close to the true state over time and this distance is quantified by the l_2 error over time and more popular metric such as RMSE. This naturally leads to the question of understanding the numerical computation LVs and their stability with respect to errors in their underlying orbits for the dynamical system at question. Above results will also let us conclude under what conditions can one use the best estimates of the state over time as the proxy of the true state whose quantification depends on the errors in the state estimate. Understanding the nature of error in the best estimate and their impact of computing these LVs will lead to discussions on further application in forecasting.

3 Methodology

3.1 Ensemble Kalman Filters

Introduced in the data assimilation community by G. Evensen Evensen (2003), ensemble kalman filters are a Monte-Carlo approach of approximation of the original Kalman filter which is a sequential state estimation algorithm Cohn (1997) based on the assumption of linearity and gaussianity of the probability distributions involved to recursively compute from the prior distribution, the posterior distribution incorporating the latest observation by simply updating mean and covariance. The filter employs ensemble representation of probability distributions which is a collection states sampled from the respective pdfs. This approach is particularly useful when the dynamics is nonlinear as each ensemble member can be integrated individually. In between the observations, each member of the ensemble is evolved in time according to the model equations. When new observation arrives, the mean and covariance required in the bayesian update is replaced by the empirical mean and covariance computed from the ensemble respectively. Localization and inflation are ad-hoc methods which makes EnKF work with small ensemble size, making it feasible for systems with large dimensions and a practical choice for operational data assimilation (Carrassi et al., 2018).

We use EnKF for twin experiment with Lorenz-96 using noisy and partial observations to assimilate the

3.2 Lyapunov Vectors from noisy trajectory

To investigate the stability of numerically calculated LVs about the reference trajectory, we add noise to the orbit by adding random samples from a standard normal distribution of covariance $\sigma^2\mathbf{I}$ at each point of the orbit. We refer to the obtained trajectory as perturbed orbit for the respective σ value in future discussions.

The calculation of BLVs over any interval requires only the forward transient which is required for the convergence of the BLVs. When using Ginelli's algorithm to compute CLV along a trajectory, we need both a forward and a backward transient interval in addition to the interval over which they are calculated, the details can be found in the paper (Ginelli et al. (2013)). We then compute all the BLVs and CLVs about the true and the perturbed trajectories over a common interval using same forward and backward transient intervals.

3.3 Lyapunov Vectors from the analysis trajectory

The analysis mean obtained from a given filter is the best estimate which takes into account all the previous observations and the current observations at that time. The sequence of analysis mean as states obtained over time is not a dynamical trajectory of the model but at any time t, the analysis mean lies close to the true state points in some truncated state space. Hence we propose to use the analysis mean of the above filter as a proxy for the true state of the system which can be used to obtain the lyapunov vectors along the trajectory the system is evolving . The above notion of sensitivity to the noise in base trajectory underlying allows us to think of the analysis mean trajectory as a perturbed trajectory obtained from the true trajectory where the perturbations follow some distribution. This analysis mean is then used for evaluating the model jacobian, a crucial step in the calculation of the LVs.

4 Experimental setting

4.1 Models

Since their first introduction by Ed. Lorenz who used them to understand the predictability of the atmospheric convection, Lorenz-63 Lorenz (1963) and Lorenz-96 Lorenz (1995) have been extensively studied and have led to understanding and development of the chaos and it's properties. Lorenz-63 model was the one of the first low-dimensional model where the emergence of chaos in low-dimensions was studied. The following set of ode describe the evolution of the system in time:

$$\frac{dx}{dt} = \sigma(y-x), \frac{dy}{dt} = x(\rho - z) - y, \frac{dz}{dt} = xy - \beta z \tag{1}$$

where, σ, ρ and β are three fixed parameters. For the specific value of $(\sigma, \rho, \beta) = (10, 28, 8/3)$, the system exhibits chaos with and has a strange attractor.

Lorenz-96, a 40-dimensional nonlinear, dissipative and a constant external forcing terms which mimics the dynamics of a meteorological scalar variable along the latitude . The model is given by the evolution of a set of n-ordinary differential equations given below:

$$\frac{dX_k}{dt} = X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F \tag{2}$$

where X_k is the k^{th} component of the n-dimensional state. For different values of F, the system is stable, weakly or strongly chaotic. It's extensive chaotic properties for different regimes of forcing and different dimension has been studied in Karimi and Paul (2010). For the specific value of forcing $F = 8$, it is a chaotic system with 13 positive lyapunov exponents and has Kalpan-Yorke dimension which equals 28.4. Due to it's small size which makes it computationally cheap, it has served as a model in twin experiments used to benchmark performance of many data assimilation algorithms before being applied to very large scale atmospheric models.

4.2 Calculation of Lyapunov vectors around a perturbed trajectory

For L63, we start from a random initial condition and integrate for a long transient time to reach the attractor. We then choose this point on the attractor as the initial condition for the true orbit which is generated by numerically integrating the ODE for a total time T=100 with $\delta t = 0.005$ using Runge-Kutta 4th order scheme and store the state every 0.01 s. We repeat the same for L96 with T=500, $\delta t = 0.01$ and saving the state every 0.05 s.

The long trajectory is divided into three regions, the forward transient T_{fr}, the sampling interval T_{si} and the backward transient T_b. For L63, we chose all the three intervals to be equal to 100. To study the dimension dependence of sensivity in L96, which had extensive chaotic property, we choose d=10, 20 and 40. For Lorenz-96 dim-40, transient intervals of 200 was found to be sufficient for the forward gram-schmidt vectors to converge to BLVs at the end of T_{fr}. Hence we fix this choice for all T_{fr}, T_b and the sampling interval T_{si} for all d.

For the choice of initial perturbations, we use the standard orthonormal vectors which are then integrated forward in time using the tangent liner equations. As the trajectory is fixed, we only evolve the perturbations using RK-4 scheme with step size=0.0. In between two consecutive points on the orbit, the state about which the jacobian J(x)is evaluated at is constant. This choice was made keeping in mind that when we compute these vectors from the perturbed trajectory, evolving the perturbed states will diverge quickly from the true trajectory. Once the frorward transient is over, we store the BLVs. The CLVs are expressed in the basis of BLVs, and the respective coefficient matrix is obtained after the backward transient is over. These coefficient matrices are then computed backwards in time via the relation below:

$$C_n = R_n^{-1} C_{n+1} \tag{3}$$

Thus, we have two sets of matrices stored, one contains the BLVs at the time instants in the sampling interval, and the other set of matrices is the co-efficient matrices, from which we get the CLVs at any time n using the following relation :

$$V_n = G_n C_n \tag{4}$$

We plot the angle between the corresponding LVs of the true trajectory and the perturbed one for different values σ from 0.1 to 0.5 in steps of 0.1. We plot the cosines of the angle between the n^{th} LV from the perturbed and the true trajectory respectively.

We use the analysis mean trajectory after a certain transient time so that the L2-error becomes bounded after which the analysis mean and trajectory is nearby for the analysis mean.

5 Results

5.1 Sensitivity for L63

We plot the three CLVs over the whole attractor in the XZ-plane. We plot the angle between the LVs about the true trajectory and the perturbed trajectory as a function of σ, the noise strength for different lyapunov vectors.

5.2 Sensitivity for L96

We then reach a limit when the LVs become poorer with increasing noise strength. Here we do not aim to filter the noise, the goal is to understand how nearby points can still be used to estimate the LVs. Noise is bounded which means the trajectory point and the noisy one distance is bounded on an average. This is relevant to the analysis state as it is near to the truth and does not constitute a dynamical trajectory of the model.

5.3 LVs obtained from the filter analysis mean

The role of CLVs along the analysis state is of potential utility in improving data assimilation algorithm. If a numerical filter performs well, then the analysis mean is sufficiently near the true state under L_2-norm. The jacobian of the whole model evaluated at the analysis is close to the jacobian evaluated at the state in terms of spectral features and are also close under the well known matrix Frobenius-norm. The analysis-mean error over time can be thought of as a perturbed orbit whose statistics can be calculated from these errors over time. Although the exact error statistics cannot be obtained in real application as one does know the underlying true orbit, we can use another important object called ensemble variance which is of the same order as the error for a reliable data assimilaton system (cite). Results - Insights of Principal angles and a few plots - Lyapunov spectrum for the cases and the Kalpan-yorke dimension. - The distribution of errors in these angles and over time.

6 Conclusions

We test the sensitivity of two types of lyapunov vectors to the added perturbations in the orbit. Secondly, we showed that with partial and noisy observations, the CLVs computed by using the analysis as a proxy for the true state is well suited for real applications as the true trajectory is both unknown and unknowable but data assimilation provides a way to approximate it using partial observations combined with the model. This could pave way for understanding and formulating their possible applications to numerical weather forecasting. We have used observation gap which is equivalent to 6 hours of the operational data assimilation centers compared to the luapunov time scale of the system. Observation frequencey higher than this will produce better estimates. The principal angle analysis also shows a promise of capturing the unstable subspaces which can be obtained from the perturbed orbits or the analysis trajectory for use in AUS like algorithms for EnKF. The effect of model errors such as truncation errors when the model is a PDE is another interesting direction which can be undertaken in future work.

Also the nature of perturbations matter even when the rmse of two different noisy trajectory is similar as seen in the two different cases of covariance matrices. Applying such analysis to high-dimensional PDE models with multiscale dynamics and spatial structures is a possible direction for future work.

Code availability. TEXT

Data availability. TEXT

Code and data availability. TEXT

Sample availability. TEXT

Video supplement. TEXT

Appendix A

A1

Author contributions. TEXT

Competing interests. TEXT

Disclaimer. TEXT

Acknowledgements. TEXT

References

Benettin, G., G. L. G. A.: Lyapunov Characteristic Exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. Part 1: Numerical application, Meccanica 15.1, pp. 21–30, https://doi.org/https://doi.org/10.1007/BF02128237, 1980.

Bocquet, M., Gurumoorthy, K. S., Apte, A., Carrassi, A., Grudzien, C., and Jones, C. K. R. T.: Degenerate Kalman Filter Error Covariances and Their Convergence onto the Unstable Subspace, SIAM/ASA Journal on Uncertainty Quantification, 5, 304–333, https://doi.org/10.1137/16M1068712, 2017.

Brugnago, E. L., Gallas, J. A. C., and Beims, M. W.: Predicting regime changes and durations in Lorenz's atmospheric convection model, Chaos: An Interdisciplinary Journal of Nonlinear Science, 30, 103 109, https://doi.org/10.1063/5.0013253, 2020.

Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G.: Data assimilation in the geosciences: An overview of methods, issues, and perspectives, WIREs Climate Change, 9, e535, https://doi.org/https://doi.org/10.1002/wcc.535, 2018.

Cohn, S. E.: An introduction to estimation theory, J. Met. Soc. Japan, 75, 257–288, 1997.

Eckmann, J. P. and Ruelle, D.: Ergodic theory of chaos and strange attractors, Rev. Mod. Phys., 57, 617–656, https://doi.org/10.1103/RevModPhys.57.617, 1985.

Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, Ocean Dynamics, https://doi.org/10.1007/s10236-003-0036-9, 2003.

Ginelli, F., Chaté, H., Livi, R., and Politi, A.: Covariant Lyapunov vectors, Journal of Physics A: Mathematical and Theoretical, 46, 254 005, https://doi.org/10.1088/1751-8113/46/25/254005, 2013.

Gurumoorthy, K. S., Grudzien, C., Apte, A., Carrassi, A., and Jones, C. K. R. T.: Rank Deficiency of Kalman Error Covariance Matrices in Linear Time-Varying System With Deterministic Evolution, SIAM Journal on Control and Optimization, 55, 741–759, https://doi.org/10.1137/15M1025839, 2017.

Karimi, A. and Paul, M. R.: Extensive chaos in the Lorenz-96 model, Chaos: An Interdisciplinary Journal of Nonlinear Science, 20, 043 105, https://doi.org/10.1063/1.3496397, 2010.

Lorenz, E. N.: Deterministic Nonperiodic Flow, Journal of Atmospheric Sciences, 20, 130 – 141, https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2, 1963.

Lorenz, E. N.: Predictability: a problem partly solved, in: Seminar on Predictability I, pp. 1–18, ECMWF, Reading UK, 1995.

Noethen, F.: Computing covariant Lyapunov vectors: A convergence analysis of Ginelli's algorithm, 2019.

Palatella, L., Carrassi, A., and Trevisan, A.: Lyapunov vectors and assimilation in the unstable subspace: theory and applications, Journal of Physics A: Mathematical and Theoretical, 46, 254 020, https://doi.org/10.1088/1751-8113/46/25/254020, 2013.

Pazó, D., Szendro, I. G., López, J. M., and Rodríguez, M. A.: Structure of characteristic Lyapunov vectors in spatiotemporal chaos, Phys. Rev. E, 78, 016 209, https://doi.org/10.1103/PhysRevE.78.016209, 2008.

Trevisan, A., D'Isidoro, M., and Talagrand, O.: Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension, Quarterly Journal of the Royal Meteorological Society, 136, 487–496, https://doi.org/https://doi.org/10.1002/qj.571, 2010.

Vannitsem, S. and Lucarini, V.: 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, 49, 224 001, https://doi.org/10.1088/1751-8113/49/22/224001, 2016.

Wolfe, C. L. and Samelson, R. M.: An efficient method for recovering Lyapunov vectors from singular vectors, Tellus A: Dynamic Meteorology and Oceanography, 59, 355–366, https://doi.org/10.1111/j.1600-0870.2007.00234.x, 2007.

Yang, H.-l., Takeuchi, K. A., Ginelli, F., Chaté, H., and Radons, G.: Hyperbolicity and the Effective Dimension of Spatially Extended Dissipative Systems, Phys. Rev. Lett., 102, 074 102, https://doi.org/10.1103/PhysRevLett.102.074102, 2009.