Ensemble Methods

Brian Hunt
10 June 2013
Goals of Lecture

• Describe a mathematical framework for ensemble methods to estimate Lyapunov exponents/vectors of a dynamical system, and/or to perform data assimilation, without explicitly linearizing the dynamics.

• Discuss work with Cecilia González-Tokman (Physica D 2012) proving that these methods “work” in an appropriate limit for hyperbolic dynamical systems.
Motivation

• For high-dimensional systems, computing the derivative of the system can be very time-consuming.
• An ensemble of nearby trajectories provides discrete information about the derivative.
• Ensemble methods can treat the system as a “black box”.
Lyapunov Exponents

• Given a trajectory of a dynamical system, the “tangent linear model” (TLM) describes the evolution of infinitesimal perturbations from that trajectory.

• Lyapunov exponents/vectors (Oseledec 1968) correspond to asymptotically exponential solutions of the TLM.

• Chaos: at least one positive Lyapunov exponent.
Given a forecast model for a physical system and an ongoing time series of observations, data assimilation is an iterative procedure to:

- Synchronize the model state with the physical state, and thereby…
- Estimate the current state of the system based on current and past observations.
Data Assimilation Cycle

• Run a weather forecast model.
• Gather atmospheric observations over a 6-hour time interval.
• Adjust the 6-hour forecast state to better fit the observations.
• Use the adjusted model state as the initial conditions for a new forecast.
• Repeat this cycle every 6 hours.
Notation and Terminology

• “Forecast model”: a discrete-time dynamical system:
  \[ x_n = f(x_{n-1}), \quad x \in \mathbb{R}^m \]

• “\( \delta \)-pseudotrajectory”: \( \{x_n\} \) for which
  \[ |x_n - f(x_{n-1})| \leq \delta \]

• “Background ensemble”: \( \{x_n^{b(i)}\} \)

• “Analysis ensemble”: \( \{x_n^{a(i)}\} \)
Ensemble Methods

• Forecast step:
  \[ x_n^{b(i)} = f(x_n^{a(i)}), \quad i = 1, 2, \ldots, k \]

• Adjustment step:
  \[ \{x_n^{a(i)}\} = g(\{x_n^{b(i)}\}, y_n), \quad y \in \mathbb{R}^\ell \]

• Preserve “ensemble space”:
  \[
  \bar{x}_n^a + \text{Span}\{x_n^{a(i)} - \bar{x}_n^a\} = \bar{x}_n^b + \text{Span}\{x_n^{b(i)} - \bar{x}_n^b\}
  \]
Example 1: Breeding

- Adjustment step:

\[
\begin{align*}
x_n^{a(1)} &= x_n^{b(1)} \\
x_n^{a(2)} &= x_n^{b(1)} + \beta \left( \frac{x_n^{b(2)} - x_n^{b(1)}}{|x_n^{b(2)} - x_n^{b(1)}|} \right)
\end{align*}
\]
Uses for Breeding

- With small $\beta$, approximate leading Lyapunov exponent/vector.
- With $\beta$ representing the size of uncertainty in initial condition, assess forecast uncertainty (Toth & Kalnay, 1993).
Ensemble Data Assimilation

• Given: observations \{y_n\} and a “forward operator” \( h \) such that
\[
y_n = h(x_n^t) + \varepsilon_n
\]
where the “truth” \( \{x_n^t\} \) is a pseudotrajectory and the “error” \( \varepsilon_n \) is usually small.

• Goal: design the ensemble adjustment operator \( g \) so that the ensemble approximates the truth well.
Ensemble Kalman Filtering

• Introduced by G. Evensen (1994).
• Formulation here based on LETKF (Hunt et al. 2007), drawing on LEKF (Ott et al. 2004) and ETKF (Bishop et al. 2002).
Ensemble Kalman Filter

• Assume (pretend) $\varepsilon_n \sim N(0,R)$ i.i.d.

• Consider $\underline{x}_n^b$ to represent the “most likely” true state given past data; $\underline{x}_n^a$ likewise but given current data too.

• Consider each ensemble to represent a Gaussian distribution with the same (sample) mean and covariance.
Ensemble Kalman Filter, cont.

- Analysis (posterior) distribution determined by Bayes’ rule from the background (prior) and observation error distributions, linearizing $h$ in ensemble space.
- Qualitatively, the adjustment step moves the ensemble toward the background members that best match the data and reduces its covariance (new information $\rightarrow$ less uncertainty).
Ensemble Kalman Filter, cont.

- Formally (square brackets → form matrix):

  \[
  \bar{x}_n^a = \bar{x}_n^b + [x_n^{b(i)} - \bar{x}_n^b]L_n(y_n - h(x_n^b)),
  \]
  \[
  [x_n^{a(i)} - \bar{x}_n^a] = [x_n^{b(i)} - \bar{x}_n^b]T_n,
  \]
  \[
  L_n = L(\{h(x_n^{b(i)})\}, R),
  \]
  \[
  T_n = T(\{h(x_n^{b(i)})\}, R)
  \]

- Remark: Breeding can also be formulated this way with appropriate \( y, h, L, T \).
ETKF specifics

• Use

\[
L_n = T_n^2 Y_n^T \left( (k-1)R \right)^{-1},
\]
\[
T_n = (I + Y_n^T \left( (k-1)R \right)^{-1} Y_n)^{-1/2}
\]

where

\[
Y_n = [h(x_n^{b(i)}) - \bar{h}(x_n^b)]
\]

• Among all \(T_n\) that give the correct analysis covariance, this minimizes distance from background to analysis ensemble.
Takens Embedding Theorem

If $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is the time 1 map of a $C^2$ flow with no orbits of integer period up to $2m+1$, and all of whose fixed points have simple eigenvalues different from 1, then for generic $C^2 h : \mathbb{R}^m \rightarrow \mathbb{R}$, the map

$$x \rightarrow (h(x), h(f^{-1}(x)), \ldots, h(f^{-2m}(x)))$$

is an embedding (one-to-one and its derivative has full rank everywhere).
Embedding Theorem, cont.

- True for diffeomorphisms more generally, and for “prevalent” $h$.
- For a $d$-dimensional attractor, the number of observations only has to exceed $2d$ (Sauer, Yorke, Casdagli 1991; following Takens 1981).
- Attractor: a compact invariant set that attracts nearby initial conditions.
Hyperbolicity

We say an attractor $A$ of a $C^1$ diffeomorphism $f : \mathbb{R}^m \to \mathbb{R}^m$ is hyperbolic if $\exists C > 0$, $\lambda > 1 > \mu$, and $\forall x \in A$, $\exists$ subspaces $E^+(x)$ and $E^-(x)$ s.t. $Df E^\pm(x) = E^\pm(f(x))$ and $\forall v \in E^+(x)$ and $n > 0$ we have $|Df^n(x)v| \geq C\lambda^n |v|$, while $\forall v \in E^-(x)$ and $n > 0$ we have $|Df^n(x)v| \leq C^{-1}\mu^n |v|$.
Proposition: Let $f$ and $h$ be as in Takens’ Theorem, and $A$ be a hyperbolic attractor w/ $<k$ unstable directions. Then $\exists C, \delta_0 > 0$ s.t. if $\delta \leq \delta_0$, $\{x_n^t\}$ is a $\delta$-pseudotrajectory, and $|\varepsilon_n| \leq \delta$, then $k$-member ETKF has the following property. For an open set of initial ensembles, the ensemble stays within $C\delta$ of the truth.

The ensemble spread in the unstable directions stays $\geq C^{-1}\delta$. 
Lyapunov Exponents from ETKF

- Recall: \[ [x_n^{a(i)} - \bar{x}_n^a] = [x_n^{b(i)} - \bar{x}_n^b]T_n \]

- Corollary: If the ensemble covariance remains bounded (above and below), the positive Lyapunov exponents/vectors of the attractor can be estimated to order \( \delta \) from the matrices \( T_n \) (we proved this for the largest Lyapunov exponent).

- Caveat: for high-dimensional systems, we can only practically estimate finite-time Lyapunov exponents/vectors.
Remarks

• For weather models, we use ensembles that are smaller than the global number of unstable directions. This works only because in LEKF/LETKF we use “localization”: assimilate in local regions.

• With similar hypotheses, we should be able to prove that for $\delta$ sufficiently small, for generic initial perturbations, breeding approximates the largest Lyapunov exponent/vector to within order $\delta$. 
Conclusions

• Ensemble methods provide a discrete analogue to algorithms that use the derivative of a dynamical system (e.g., standard methods for computing Lyapunov exponents, Extended Kalman Filter, 4D-Var).

• We can prove convergence results for ensemble methods in hyperbolic systems…and beyond??