Projected Data Assimilation

Erik S. Van Vleck
Department of Mathematics
University of Kansas
erikvv@ku.edu
people.ku.edu/$\sim$erikvv

SIAM Annual Meeting
Data Assimilation and Nonlinear Filtering

July 11, 2017

Thanks

Old Stuff (Numerical Shadowing):

Shui-Nee Chow (Atlanta).

Old Stuff (Computation of Lyapunov Exponents):

Mohamed Badawy (Austin), Luca Dieci (Atlanta), Cinzia Elia (Bari).

New Stuff (Projected DA):

Bart de Leeuw (Amsterdam), Svetlana Dubinkina (Amsterdam), Jason Frank (Utrecht), Andrew Steyer (Albuquerque), Xuemin Tu (Lawrence).

Project Goals

The goals of the project (work in progress) are to combine:

- the use of sophisticated models of land-surface, radiation-convection, ground-water flow, ...
- data collected from from different sources,
- computational dynamical systems, data assimilation, parameter estimation, and uncertainty quantification techniques.

to investigate problems related to climatological change and obtain improved understanding of uncertainties.

Data Assimilation

For $n = 0, 1, ..., N-1$,

Data Model $y_{n+1} = H(u_{n+1}) + \eta_{n+1}$

State Space Model $u_{n+1} = F(u_n) + \xi_n$

Here y_n are the observations, η_n observational noise, and ξ_n the model noise.

Idea is to use the observations to determine solution, i.e., u_0.

Common techniques (based on Bayes' Rule): Variations on ensemble Kalman filter, 4DVar, proposal density methods, particle filters, etc..

Consider a nonlinear evolution equation (solution operator of a model)

$$u_{n+1} = F_n(u_n; \alpha), \quad n = 0, 1, ..., N-1$$

where

- u_n are the state variables at time n,
- α are adjustable model parameters, e.g., global in time.

Write $u_n = u_n^{(0)} + \delta_n$.

If we can decompose the time dependent tangent space into slow variables and fast variables, then we write $\delta_n = P_n \delta_n + (I - P_n)\delta_n$.

Rewrite original nonlinear evolution approximately as two subsystems ...

For $k = 0, 1, 2, ...$

P System:

$$u_{n+1}^{(k)} + P_{n+1}\delta_{n+1} = P_{n+1}F_n(u_n^{(k)} + P_n\delta_n),$$
$$u_{n+1}^{(k+\frac{1}{2})} = u_n^{(k)} + P_n\delta_n, \;\; n = 0, 1, ..., N-1,$$

I-P System:

$$u_{n+1}^{(k+\frac{1}{2})} + (I - P_{n+1})\delta_{n+1} = (I - P_{n+1})F_n(u_n^{(k+\frac{1}{2})} + (I - P_n)\delta_n),$$
$$u_n^{(k+1)} = u_n^{(k+\frac{1}{2})} + (I - P_n)\delta_n, \;\; n = 0, 1, ..., N-1.$$

Roughly speaking the first subsystem contains the slow variables (positive, zero, and slightly negative Lyapunov exponents) and the second subsystem contains the fast variables (strongly negative Lyapunov exponents).

Some questions ...

Q1: Computing time dependent projections P_n and $(I - P_n)$?

Q2: How to determine $u^{(0)}$, its uncertainties, etc.?

Q3: How to take advantage of such a slow/fast splitting?

Notes:

- Assimilation in the Unstable Subspace (AUS) [Trevisan, D'Isidoro, Talagrand '10 Q.J.R. Meteorol. Soc., Palatella, Carrassi, Trevisan '13 J. Phys A, ...]
- Error analysis in DA for hyperbolic system [González-Tokman, Hunt '13 Phys D]
- Convergence of covariances matrices in unstable subspace [Bosquet etal. '17 SIAM UQ]

Importance of observations rich in unstable subspace

Time Dependent Stability Theory

- Lyapunov exponents play the role of the real parts of eigenvalues in time dependent stability theory,
- Computational techniques involve performing a change of variables to effectively extract the Lyapunov exponents and corresponding Lyapunov vectors,
- For review of recent developments on computation of Lyapunov exponents, see

  Dieci & VV (2015) Encyclopedia of Applied and Computational Mathematics, 834–838.

  VV (2015) Festschrift chapter in honor of Volker Mehrmann in Numerical algebra, matrix theory, differential-algebraic equations and control theory, 299-318.

Forming time dependent projections

Discrete QR algorithm for determining Lyapunov exponents, local in time stability information, etc.:

For $Q_0 \in \boldsymbol{R}^{m\times p}$ random such that $Q_0^T Q_0 = I$,

$$Q_{n+1}R_n = F'(u_n)Q_n, \;\; n = 0, 1, \ldots$$

where $Q_{n+1}^T Q_{n+1} = I$ and R_n is upper triangular with positive diagonal elements.

For high dimensional models or when the tangent linear model is not explicitly known, finite difference approximation

$$F'(u_n)Q_n \approx \frac{1}{\epsilon}[F(u_n + \epsilon Q_n) - F(u_n)], \;\; \epsilon \approx \sqrt{\epsilon_M}\|F(u_n)\|$$

Forming time dependent projections

Windowed Growth/Decay Rates:

$$\lambda_j(n, M) = \frac{1}{M} \sum_{k=n}^{n+M-1} \log(R_k(j, j))$$

The columns of Q_n form an orthonormal basis at time n for initial conditions with growth/decay rates $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$

Orthogonal Projections:

$$P_n = Q_n Q_n^T, \;\; (I - P_n) = I - Q_n Q_n^T$$

Forming time dependent projections

- Under assumptions (generic on half line $t > 0$) on the continuity of Lyapunov exponents with respect to perturbations in $F'(x_n)$:
  - the columns of Q_n form an orthonormal basis for the Lyapunov vectors associated with first p Lyapunov exponents at time n [Dieci, VV '02, '07, Dieci, Elia, VV '10, '11],
  - the Q_n are robust with good long time global error properties [Dieci, VV '05, '06, '08, VV '10, Badawy, VV '12].

Initial approximation $u^{(0)}$

Data assimilation schemes typically make use of a prior/backgound initial condition u_0^b
For example,

- for Kalman filter techniques u_0^b is the prior mean,
- for variational techniques such as 4DVar u_0^b serves as initial guess for optimization.

Kalman filter and variational techniques can be used to determine $u^{(0)} \equiv \{u_n^{(0)}\}_{n=0}^{N}$.

Initial approximation $u^{(0)}$

Another approach (essentially an insertion technique) is to combine u_0^b with available observations:

Given observations y_n, $n = 0, 1, ..., N$ that are a subset of the state space variables, let

$$u_0^{(0)} = y_0 \cup H^{\perp}(u_0^b),$$

$$\tilde{u}_{n+1}^{(0)} = F(u_n^{(0)}), \;\; u_{n+1}^{(0)} = y_{n+1} \cup H^{\perp}(\tilde{u}_{n+1}^{(0)}), n = 0, 1, ...$$

Slow/Fast Splitting and Techniques for P and $I - P$

Using the framework of a slow/fast splitting we

- may employ different DA/parameter estimation techniques in each subsystem,
- obtain an explicit representions for the time dependent unstable subspace.

Interested in systems that are hyperbolic in flavor (finite number of positive Lyapunov exponents, few zero Lyapunov exponents, potentially many negative Lyapunov exponents).

Slow/Fast Splitting and Techniques for P and $I - P$

Techniques for P System:

- Particle filters that are effective for low dimensional problem, a new class of techniques based upon residual correction (Pseudo Orbit DA (PDA) [Du & Smith I & II, '14 J. Atmos. Sci. 2014], shadowing refinement [Grebogi, Hammel, Yorke, and Sauer, Phys Rev Lett (1990)]), ...

Techniques I-P System:

- Techniques such as ETKF, LETKF, 4DVar (essentially a shooting method starting from u_0^b, trying to match observations, basin of attraction shrinks in the presence of positive Lyapunov exponents).

Shadowing Refinement and Parameter Estimation

We've developed "interval sequential" (all "observations" over subintervals are employed simultaneously) data assimilation/parameter estimation techniques based upon

- Shadowing refinement and parameter estimation for P problem,
- Insertion synchronization or ETKF for $I - P$ problem.

We first describe shadowing refinement and parameter estimation w/o projection ($P = I$):

Given $u^{(0)}$ we solve $G(u) = 0$ where

$$(G(u); \alpha)_n = u_{n+1} - F(u_n; \alpha), \quad n = 0, 1, \ldots, N - 1.$$

Shadowing Refinement and Parameter Estimation

This is N vector residuals with $N+1$ vector unknowns plus q parameters so underdetermined.

Thus, we employ Gauss-Newton where linear systems are solved using pseudo inverse so use minimum two norm solution as update.

If no parameter estimation ($q=0$) then for $L \approx D_u G(u)$ pseudo inverse is (LL^T block tridiagonal)

$$L^+ = L^T (LL^T)^{-1}$$

$$\delta^{(k)} = -L^+ G(u^{(k)}), \;\; u^{(k+1)} = u^{(k)} + \delta^{(k)}$$

Parameter Estimation

With no projection and with parameter estimation ($q > 0$),

$$L = G'(u; \boldsymbol{\alpha}) = [D_u G(u; \boldsymbol{\alpha}) | D_{\boldsymbol{\alpha}} G(u; \boldsymbol{\alpha})]$$

and

$$LL^T = D_u G(u; \boldsymbol{\alpha}) D_u G(u; \boldsymbol{\alpha})^T + D_{\boldsymbol{\alpha}} G(u; \boldsymbol{\alpha}) D_{\boldsymbol{\alpha}} G(u; \boldsymbol{\alpha})^T$$

which is a rank q perturbation of the block tridiagonal $D_u G(u; \boldsymbol{\alpha}) D_u G(u; \boldsymbol{\alpha})^T$

Using Sherman-Morrison-Woodbury formulas, linear systems $LL^T x = b$ can be solved with $q + 1$ block tridiagonal solves.

In Projected Space

Time dependent rank p orthogonal projections $P_n = Q_n Q_n^T$ formed using

$$Q_{n+1} R_n = D_{u_n} F(u_n; \boldsymbol{\alpha}) Q_n$$

The matrix L (with $m \times m$ blocks) replaced by $\tilde{L}$ (with $p \times p$ blocks)

$$(L\delta)_n = \delta_{n+1} - D_{u_n} F(u_n; \boldsymbol{\alpha}) \delta_n$$

and

$$(\tilde{L} z)_n = z_{n+1} - R_n z_n, \;\; \delta_n = Q_n z_n.$$

Extension to optimizing parameters in projected space (still rank q perturbation of block tridiagonal).

Simultaneous state space/parameter space estimation without introducing additional neutral modes.

Example: Lorenz '96 Model ($N = 40, F = 8$)

$$\dot{u}_k = (u_{k+1} - u_{k-2})u_{k-1} - u_k + F \ , \ k = 0, 1, \ldots, N-1 \ , \ (\mathrm{mod}\ N)$$

- $\dot{u} = -Iu + N(u)$
- 13 positive Lyapunov exponents, Lyapunov dimension ≈ 28.
- Start from a high precision numerically generated solution, noisy orbit using matlab command `2*(2*rand()-1)` .

p	$\frac{1}{N}\sum_{i=1}^{N} \|u_i - y_i^o\|_2^2$
5	125.1
10	58.5
13	32.2
20	30.4
30	27.7
40	26.8

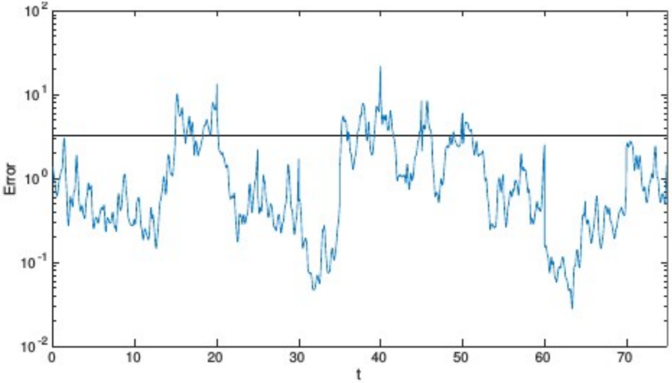
Simultaneous State Space and Parameter Estimation

Example: Lorenz '95 (Noisy = Truth($F = 8$) + $\epsilon \cdot \texttt{randn()}$)

F_0	ϵ	F_∞
2	0	7.9996
4	0	7.9997
12	0	8.0003
2	2	8.53
2	1	8.04
2	1/2	8.02
2	1/4	7.97
4	2	8.52
4	1	8.04
4	1/2	8.02
4	1/4	7.97
12	2	8.51
12	1	8.04
12	1/2	8.02
12	1/4	7.97

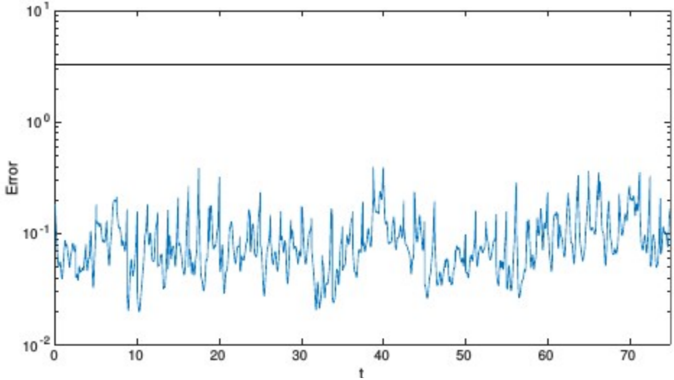
Interval Sequential

Example: Lorenz '95 (Noisy = Truth($F = 8$) + 0.3·`randn()`)



RMSE with window lengths of 5 and $p = 15$.

Interval Sequential



RMSE with window lengths of 1.25 and $p = 15$.

Continuity between subintervals imposed in the strongly stable subspace (v_T terminal value from previous window):

$$(I - P_0)\delta_0 = (I - P_0)[v_T - u_0^{(k+1/2)}]$$

Applications

Working on several different applications

- Noah-MP: a land surface model combined with satellite and tower data,
- ModFlow: A ground-water flow model together with hydraulic conductivity and pressure head data.
- RadCon: A single column radiation-convection model of K. Emanuel together with TOGA COARE data,

These models all have relevant configurations in which the number of positive and near zero Lyapunov exponents is relatively small as compared to the problem dimension.

Conclusions and Future Work

- Expansion in terms of Lyapunov vectors to determine splitting of perturbed nonlinear models.
- The technique we have focused on is in some sense a hybrid of
  - synchronization,
  - and non-autonomous inertial manifold technique.
- Have some theoretical results for synchronization scheme for both linear non-autonomous and nonlinear models, essentially: Good approximation in unstable subspace implies good asymptotic synchronization independent of initial condition in strongly stable subspace.
- Choice of methods to employ for P and $I - P$ problems.
- Form projections, apply, e.g., Bayesian techniques to P and $I - P$ problems, combine results to obtain means, covariances, etc..