Dynamical Systems, Data Assimilation

Douglas Nychka,
National Center for Atmospheric Research
www.image.ucar.edu/~nychka

Seattle, August 2012
● Nonlinear maps
● Lorenz ’96 weather and chaos
● What is data assimilation
● Making a forecast

A 40 variable, nonlinear model, Lorenz ’96, illustrates principles of nonlinear dynamical systems.
Given a model for the atmosphere and observations what is the full state?

- **Data assimilation**: statistical techniques to combine numerical models with observations to improve estimate of the state of a system.
- Typically has a sequential aspect – observation are spaced in time.
- The system can have complicated nonlinear dynamical behavior and the state vector may be large.

**This is a problem in Bayesian statistics.**
With linear process dynamics and Gaussian distributions the methods become the Kalman Filter.

**This is like a problem in spatial statistics**
Updating geophysical fields with observations is the same as spatial prediction.
The model tuning problem

\[ x(t) = \Gamma(x(t - 1), F, \theta) \]

\( \Gamma \) is a (nonlinear) map, a model for the atmosphere.

**Weather:** \( x(t) \)

**Climate:** \( E[x(t)] \) or \( \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} x(t) \)

**What we observe:**
\( y_i = \) components of \( x(t_i) + \) error

\( y_i \) is essentially observed weather at time \( t_i \)

**The problem:**
Can we use the observations and the model \( \Gamma \) to estimate forcings and parameters, \( F \) and \( \theta \)?
An introduction to data assimilation
The basic data assimilation cycle

The problem is to estimate the state of a system $x_t$ (or a field) at different times based on incomplete and irregular observations. This is how the weather is forecast.

- Start with a forecast $x_t^f$ for the state at time $t$ (PRIOR)
- New data, $y$ comes in at time $t$ (LIKELIHOOD)
- Update $x_t^a$ in light of the data (POSTERIOR)
- Forecast ahead to time $t + 1$ using updated state.
- *Cycle repeats* – new obs come in at $t + 1$.

"a" stands for the analysis field.
Making a forecast

Γ a dynamical numerical model for the system

\[ x_{t+1} = \Gamma(x_t) \]

If \( x^a_t \) is the best estimate of the state :

\[ x^f_{t+1} = \Gamma(x^a_t) \]

The big idea:
*Obtain an ensemble of forecasts by applying \( \Gamma \) to each ensemble member.*
Bayes approach for estimating the state

*(Ignoring parameters for the moment)*

Provides a solution by translating the problem into sequential probability distributions.

**PRIOR** $[x_t]$ probability density function for the state

**LIKELIHOOD** $[y_t|x_t]$ probability density function for observations given knowledge of the state.

**POSTERIOR** $[x_t|y_t]$ ... state given ... observations

Posterior found by Bayes Theorem:

$$[x_t|y_t] \propto [y_t|x_t][x_t]$$
Making a forecast

The weather man:
\( \hat{x}_t \) is the best estimate of the current state of the atmosphere

Propagate this to time \( t + 1 \):

\[ x^f_{t+1} = \Gamma(\hat{x}_t) \]

The Bayesian:
\[ [x_t|y_t] \text{ the posterior} \]

Propagate this to time \( t + 1 \):

\[ [x^f_{t+1}|y_t] = [\Gamma(x_t)|y_t] \]
Linear and Gaussian

(Ignoring parameters for the moment)
Provides a solution by translating the problem into sequential probability distributions.

**PRIOR** \( [x_t] \) probability density function for the state

\[ x_t \sim MN(\mu_f, \Sigma_f) \]

**LIKELIHOOD** \( [y_t|x_t] \) probability density function for observations given knowledge of the state.

\[ y_t = Hx_t + MN(0, \Sigma_o) \]

\( H \) is a known matrix that connects the mean of the observations to the state.
Linear and Gaussian (continued)

Posterior found by Bayes Theorem:

\[ \text{POSTERIOR} \ [x_t|y_t] \ ... \ \text{state given} \ ... \ \text{observations} \]

\[ [x_t|y_t] \propto [y_t|x_t][x_t] \]

\[ x_t^a \sim MN(\mu_a, \Sigma_a) \]

- Mean vector and covariance given by the Kalman filter!
- This update is also just spatial statistics/ conditional normals.
- Also a solution to a minimization problem (3-d VAR)
The formula for the mean of the posterior has another form:

$$\mu_a \text{ minimizes over } u:$$

$$(y - Hu)^T \Sigma_o^{-1} (y - Hu) + (u - \mu_f)^T \Sigma_f^{-1} (u - \mu_f)$$

Fit to observations + closeness to prior
Problems with an exact Bayes approach

Recall: 
\[ [x^f_{t+1}|y_t] = [\Gamma(x_t)|y_t] \]

Problem 1:
Propagating the distribution forward to \( t+1 \) is the mother of all change-of-variables problems!

Problem 2:
Update formula for the mean

\[ \mu_a = \mu_f + \sum_f H (H^T \Sigma_f H + \Sigma_o)^{-1} (y - H \mu_f) \]

In a typical weather prediction application there are \( 10^5 \) observations at \( t \) and the state is \( 10^6 \).

\( H^T \Sigma_f H \) is \( 10^5 \times 10^5 \) – HUGE!
These are both deal breakers.
Ensemble approximation

Each distribution is represented by a random sample of the states called an \textit{ensemble}.

In place of

\[ [x_t | y_t] \rightarrow [g(x_t) | y_t] \]

propagate each ensemble member.

\[
\begin{align*}
\begin{array}{c}
\Gamma(x_{t,1}) = x_{t+1,1} \\
\Gamma(x_{t,2}) = x_{t+1,2} \\
\vdots \\
\Gamma(x_{t,M}) = x_{t+1,M}
\end{array}
\end{align*}
\]

\textbf{By elementary probability:}

\{x_{t,j}\} is a random sample from \([x_t | y_t]\) implies \{x_{t+1,j}\} will be a random sample from \([x_{t+1} | y_t]\)
Wherever a covariance matrix or mean vector appears in the posterior replace these by the sample quantities from the ensemble.

**Update formula for the mean**

\[
\mu_a = \mu_f + \Sigma_f H (H^T \Sigma_f H + \Sigma_o)^{-1} (y - H \mu_f)
\]

- This has much lower rank than the full matrices.

Jeff Anderson is able assimilate with the global NCAR atmospheric model (state vector \(10^6\)) with an 80 member ensemble.

- The covariance matrix is tapered to be a more stable estimate and inflated to make the filter stable.

- Observation vector can be sequentially assimilated one value at a time.
**Overall algorithm:**

- Ensemble of states represents prior at $t$
- Each Ensemble is updated by $y_t$ to represent a sample from posterior
- Ensemble is propagated forward in time to give new prior at $t + 1$

Similar to a particle filter but the particles are always changing and have a constant equal weight.
Jeff’s excellent Adjustment KF

Prior: \( \{ x_j^{OLD} \} \) ensemble of states

Find \( \Sigma_f \) based on the the sample covariance matrix of the prior ensemble (and also taper it) find \( \mu_f \) as sample mean.

Compute Prior obs: \( \{ \hat{y}_j^{OLD} = Hx_j^{OLD} \} \)

Compute Posterior obs:
• \( \hat{y}_j^{NEW} \) a linear transformation of \( \{ \hat{y}_j^{OLD} \} \) so that the ensemble has (sample) mean: \( H\mu_a \) and (sample) covariance matrix: \( \Sigma_o(H\Sigma_fH^T + \Sigma_o)^{-1} \)

Finding the mean is where the new observations figure in.

Transform increments to state vectors

\[
x_j^{NEW} - x_j^{OLD} = (\Sigma_fH^T)(H\Sigma_fH^T)^{-1}(\hat{y}_j^{NEW} - \hat{y}_j)
\]
Proof of concept for the Adjustment Kalman filter
The Lorenz ’96 system


A toy model for ”flow” on a ”latitude band”.

\[
\frac{dx_j}{dt} = -x_{(j-2)}x_{(j-1)} + x_{(j-1)}x_{(j+1)} - x_j + F_j
\]

\[F_1\ldots F_{30} = 6 \text{ and } F_{31}\ldots F_{40} = 12\]
A space and time diagram of L’96

Running the model with random initial conditions for 400 steps.

Hoefmuller diagram

System at time=100
The System’s “Climate”

Mean and standard deviation of state vectors.
Noisy observations

Observations: \( y_{j,t} = x_{j,t} + N(0, R) \)

State equation: \( x_t = g(x_{t-1}, F) \)

With just \( y_t \)'s, what can we say about \( x_t? \) \( F? \)

- This is a filtering problem or an inverse problem or a nonlinear time series problem.

- We want to consider methods that do not require repeated evaluations of the data.
  (e.g. simple maximum likelihood will not work.)
Some points

- Traditionally, the expectation of \( [x_t|y_t] \) is used as a prediction of the state.
- We obtain \( [x_{t+1}|y_t] \) using our knowledge of the dynamical system, \( g \).
- We can append unknown parameters to the state vector and apply the same formalism to an extended state

\[ x^* = (x, F) \]

This lets us find \( [F|y_1, \ldots, y_n] \)
Proof of concept for Lorenz ’96

Estimate $F_j$ based on just a sequence of noisy observations over time using the Adjustment Kalman Ensemble filter (80 members)

We need to add some extra variation to make the Ensemble Kalman Filter work.

$$F_{j,t+1} = F_j + \text{white noise}_t$$
Proof of concept (continued)

Truth

Estimates

D. Nychka Data Assimilation
Estimates of forcing parameters at three locations:
How well have we tuned climate?

Inference based on 80 member ensemble
We do not really understand why these approaches work!

We do not understand the exact problem we are solving!
A more interesting dynamical system
The setup

$\Gamma$: NCAR global atmospheric model with ocean and land surfaces prescribed using observations.

$y_t$: Nearly all available measurements on the atmosphere (balloons, surface, satellite and aircraft) every 6 hours.
Finding a physical model parameter

The gravity wave drag efficiency parameter in the NCAR atmospheric model.

Flow over mountains or thunderstorms (convection) produce vertical motion in the atmosphere that effects upper level flow – at global scales.

This parameter controls momentum transfer from the surface into the upper atmosphere – and its value is chosen empirically.
Results for CAM - T42.

Estimating parameters based on actual observations

Small values over ocean – as expected.
Large value near tropical mountains (New Guinea).

Parameter values compensate for other errors!
Summary

- Using approximate Bayesian approaches it may be possible to estimate parameters for large systems based on observations.

- There are many open statistical problems for assimilating data into dynamical systems.

- Challenges for estimating parameters may lie with finding informative observations and identifying parameters that can be tuned!