Dynamical Systems and Uncertainty

Eric J. Kostelich

Arizona State University
School of Mathematical & Statistical Sciences

ICTS Data Assimilation Research Program
July 6, 2011

Co-workers and Web resources

Istvan Szunyogh, Brian Hunt, Edward Ott,

Eugenia Kalnay, Jim Yorke

and many others!

Thanks to: NSF, NASA, ASU

Papers, preprints, and codes:

http://www.weatherchaos.umd.edu
http://math.la.asu.edu/∼eric

Principal papers

Preprints: www.weatherchaos.umd.edu

Initial papers:

- E. Ott *et al.*, *Tellus A* **56** (2004), 415–428
- I. Szunyogh *et al.*, *Tellus A* **57** (2005), 528–545

Refined mathematical implementation: B. R. Hunt, E. K., I. Szunyogh, Physica D **230** (2007) 112–126

Results with real data: I. Szunyogh, E.K. et al., Tellus A **60** (2008) 113–130

Outline of the lecture

- Relevant theory of ordinary differential equations
- A little bit about chaos
- Forecast ensembles and local low dimensionality

Picard's existence and uniqueness theorem

- Suppose $\mathbf{f}(\mathbf{x}, t)$ is Lipschitz continuous in a neighborhood N of $(\mathbf{x}_0, t_0)$, i.e.,

$$\|\mathbf{f}(\mathbf{x}, t) - \mathbf{f}(\mathbf{y}, t)\| \leq L\,\|\mathbf{x} - \mathbf{y}\|$$

  for some constant L whenever $\mathbf{x}, \mathbf{y} \in N$
- Then the initial value problem

$$\mathbf{x}' = \mathbf{f}(\mathbf{x}, t) \quad \text{with} \quad \mathbf{x}(t_0) = \mathbf{x}_0$$

  has a unique solution in an interval around t_0 (the size of which depends on N and $\mathbf{f}$)
- Implication:

  Perfect initial data $\Longrightarrow$ Perfect predictability

Gronwall's inequality

- Given $\mathbf{f}$, N, and L as before, and suppose that $\hat{\mathbf{x}}(t_0) = \hat{\mathbf{x}}_0$ approximates $\mathbf{x}(t_0) = \mathbf{x}_0$
- Then

$$\|\mathbf{x}(t) - \hat{\mathbf{x}}(t)\| \leq \|\mathbf{x}_0 - \hat{\mathbf{x}}_0\| \, e^{L(t-t_0)}$$

- This is the best estimate that we can expect in general
- Example:

$$x' = Lx \quad \text{with} \quad x(0) = x_0 \quad \text{and} \quad \hat{x}(0) = \hat{x}_0$$

Then

$$|x(t) - \hat{x}(t)| = |x_0 - \hat{x}_0| \, e^{Lt}$$

A hint of the difficulties

- Uncertainties in initial conditions may amplify exponentially in time!
- The details are highly equation dependent
- Example: $x' = -Lx$ has the same Lipschitz constant, but

$$|x(t) - \hat{x}(t)| = |x_0 - \hat{x}_0| \, e^{-Lt} \to 0 \quad \text{as} \quad t \to \infty$$

- Under what circumstances do uncertainties grow?

Simple case: Linear systems with constant coefficients

- Suppose $\mathbf{x}' = \mathbf{A}\mathbf{x}$ where $\mathbf{A} \in \mathbf{R}^{n \times n}$ has n distinct real eigenvalues
- The initial condition $\mathbf{x}(0) = \mathbf{x}_0$ yields the solution

$$\mathbf{x}(t) = \mathbf{x}_0 e^{\mathbf{A}t} = c_1 e^{\lambda_1 t}\mathbf{v}_1 + \cdots + c_n e^{\lambda_n t}\mathbf{v}_n$$

  where $[\mathbf{x}_0]_V = (c_1, \ldots, c_n)^{\mathrm{T}}$ in the basis of eigenvectors
- Analogous results for repeated and complex eigenvalues

Net result: Linear systems with constant coefficients

- Errors in initial conditions in the system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ grow exponentially with time whenever $\mathbf{A}$ has a positive eigenvalue (or an eigenvalue with positive real part)
- This is a global result

Harder case: Nonlinear systems

- Local result: Suppose $\mathbf{x}_0$ is a fixed point for $\mathbf{x}' = \mathbf{f}(\mathbf{x})$
- $\mathbf{x}_0$ is hyperbolic if the eigenvalues of the Jacobian matrix $\mathbf{A} = \mathbf{Df}(\mathbf{x}_0)$ are all nonzero (or have nonzero real part)
- Hartman-Grobman theorem: There exists a change of coordinates that maps solutions of $\mathbf{x}' = \mathbf{f}(\mathbf{x})$ onto solutions of the linear system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ in a neighborhood of $\mathbf{x}_0$ whenever $\mathbf{x}_0$ is hyperbolic

Basic classification of hyperbolic fixed points

Sink: All eigenvalues negative (or negative real part)

Saddle: Some eigenvalues negative and some positive (or some negative and some positive real parts)

Source: All eigenvalues positive (or positive real part)

- Sinks are stable, i.e., insensitive to small initial errors.
- Saddles and sources are unstable (sensitive)

Example: The damped nonlinear pendulum

Assume linear friction:

$$x'' + kx' + \sin x = 0 \quad \text{with} \quad k > 0$$

Define $x_1 = \text{position}(= x)$ and $x_2 = \text{velocity}(= x')$ The equivalent first-order system is

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= -kx_2 - \sin x_1 \end{aligned}$$

There are two fixed points:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \pi \\ 0 \end{pmatrix}$$

Fixed point analysis

The linearized equation about each fixed point $\mathbf{p}$ is

$$\mathbf{x}' = \mathbf{Df}(\mathbf{p})\,\mathbf{x} = \begin{pmatrix} 0 & 1 \\ -\cos x_1 & -k \end{pmatrix} \mathbf{x}$$

At $\mathbf{p} = (0,0)$: $\mathbf{Df}(0,0) = \begin{pmatrix} 0 & 1 \\ -1 & -k \end{pmatrix}$ with eigenvalues

$$\lambda_{\pm} = \frac{-k \pm \sqrt{k^2 - 4}}{2}$$

so $\lambda_{\pm} < 0$ (or $\mathrm{Re}\,\lambda_{\pm} < 0$). So $(0,0)$ is a sink (stable)

Fixed point analysis, 2

- At $\mathbf{p} = (\pi, 0)$: $\mathbf{Df}(\pi, 0) = \begin{pmatrix} 0 & 1 \\ +1 & -k \end{pmatrix}$ with eigenvalues

$$\lambda_{\pm} = \frac{-k \pm \sqrt{k^2 + 4}}{2}$$

- $\lambda_- < 0 < \lambda_+$. Hence $(\pi, 0)$ is a saddle (sensitive)
- Initial perturbations grow exponentially (at least initially)

The geometry of uncertainty

- Suppose our knowledge of the initial condition $\mathbf{x}_0$ is a "circle" of uncertainty (i.e., the underlying pdf is circularly symmetric and centered about $\mathbf{x}_0$)
- How does a dynamical system propagate the uncertainty?

Linear example: $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$

- Basic formula: matrix × circle = ellipse
- Key idea: The singular value decomposition

$$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times n}\mathbf{S}_{n\times n}\mathbf{V}^{\mathrm{T}}_{n\times n}$$

- $\mathbf{S} = \mathrm{diag}(s_1, s_2, \ldots, s_n)$ gives the singular values, which are the square roots of the eigenvalues of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$
- By convention, $s_1 \geq s_2 \geq \cdots \geq s_n \geq 0$
- If C is the unit circle, then s_i is the length of the ith axis of $\mathbf{A}C$

Example: the damped nonlinear pendulum with $k = 2$

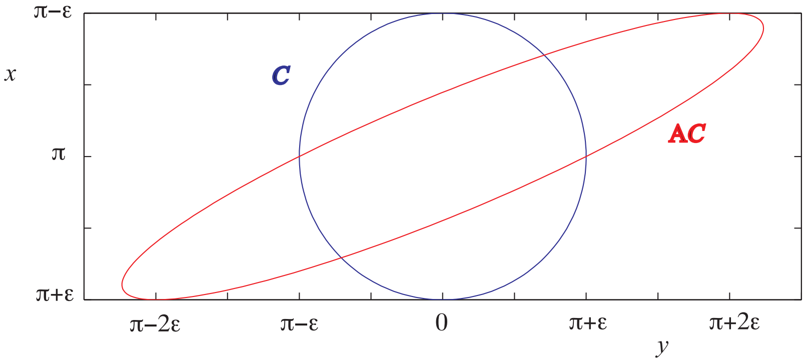
- Equation: $x'' + 2x' + \sin x = 0$
- Consider the unstable fixed point at $(\pi, 0)$
- The previous analysis shows that the eigenvalues are

$$\lambda_{\pm} = \frac{-k \pm \sqrt{k^2 + 4}}{2} = -1 \pm \sqrt{2}$$

- The Jacobian matrix is

$$\begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix}$$

Example: the damped nonlinear pendulum with $k = 2$



$$s_1 = 1 + \sqrt{2} \approx 2.4 \qquad s_2 = \sqrt{2} - 1 \approx 0.4$$

Forced, damped, nonlinear systems

- When damped nonlinear systems are forced strongly enough, they often become chaotic
- In a chaotic process, every point is a sensitive point
- Uncertainties in the initial condition of a chaotic process make it hard to predict—even if the process is deterministic

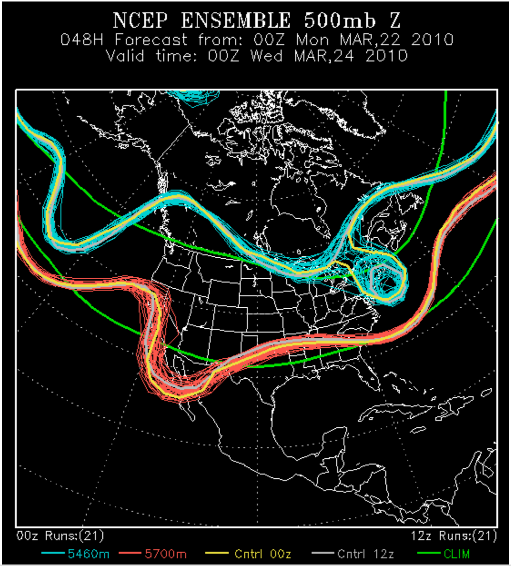
Ensemble forecasting

- How does a nonlinear model propagate a "circle" of uncertainty?
- One procedure: Given $\mathbf{x}' = \mathbf{f}(\mathbf{x})$, integrate the variational equations $\mathbf{U}' = \mathbf{Df}(\mathbf{x})\,\mathbf{U}$
- Not simple to do if $\mathbf{f}$ is big and complicated
- Simpler procedure: Integrate an ensemble of statistically equivalent initial conditions to approximate the uncertainty

Ensemble forecasting, 2

- Ensemble forecasting is a Monte Carlo method: all state vectors are treated as random variables
- Take an initial condition that represents the “best guess” (maximum likelihood estimate) of the current atmospheric state
- Choose nearby initial conditions (relative to some norm) that sample the estimated error covariance of the best guess
- Spaghetti plots are one visualization method

Spaghetti plot of a typical 72-hour forecast



Key points

- The weather is a chaotic dynamical process
- Forecast uncertainty grows exponentially over short time scales...
- ...and varies considerably in time and space
- Lorenz's estimate: The uncertainty in the global atmospheric state vector roughly doubles every 48 hours
- Places an upper bound on the predictability of the weather: $\sim$ 2 weeks

The role of data assimilation

- Without periodic corrections, the forecasts produced by a weather model would be no better than climatology
- Data assimilation combines model forecasts and atmospheric observations into an updated “best guess” of the global atmospheric state vector (and possibly its uncertainty)
- Some mathematical questions:
    - Find a useful representation of the forecast uncertainty
    - Use the available observational data efficiently
    - Estimate systematic errors (biases) in observations and forecast models
    - Where might extra observations be targeted to best advantage?

Local low dimensionality

- Most geophysical models have millions of variables
- But locally they often act like relatively low dimensional dynamical systems
- A medium-resolution weather model has ~ 3000 variables in a typical 1000×1000 km^2 synoptic region ($\sim$Texas, France)
- We can represent a forecast ensemble over such a region as a $3000 \times k$ matrix $\mathbf{X}_F$
- The singular values and vectors of $\mathbf{X}_F$ can provide excellent low-rank approximations

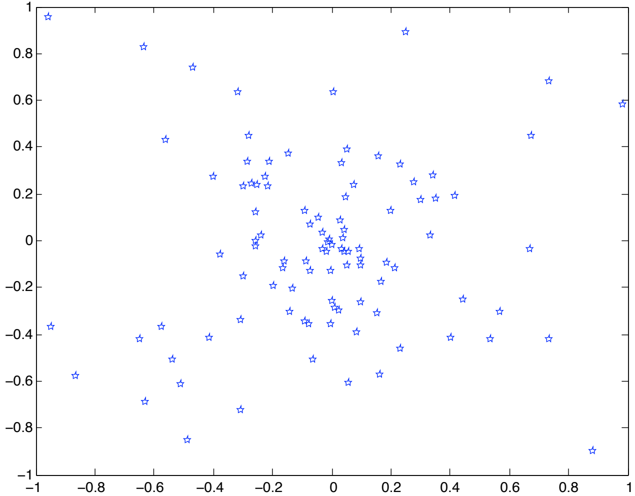
The ensemble dimension

- The ensemble dimension (E-dimension) of an $n \times k$ matrix is

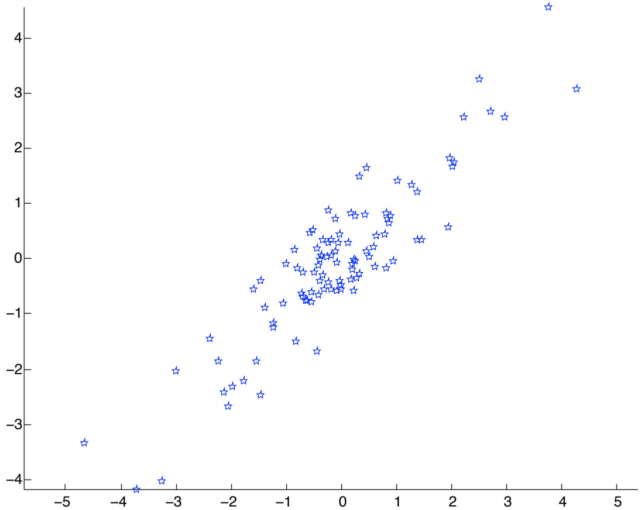
$$E \equiv \frac{(s_1 + s_2 + \cdots + s_k)^2}{s_1^2 + s_2^2 + \cdots + s_k^2}$$

- Measures the eccentricity of the "ellipse" of forecast uncertainty

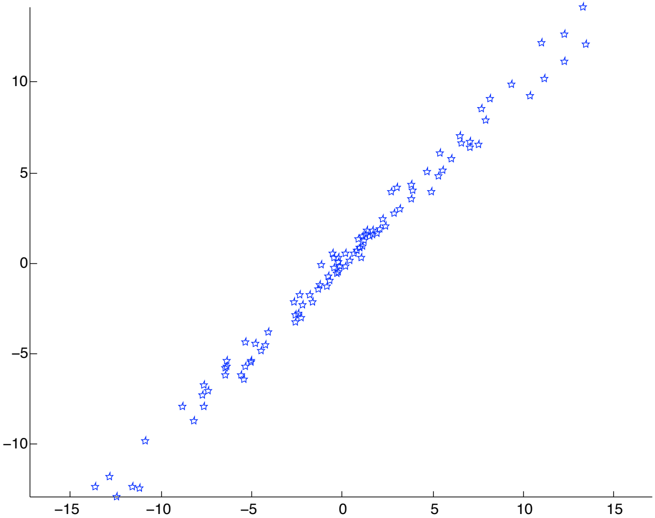
Example: $s_1 = 3.78$, $s_2 = 3.60$, $E_{\text{dim}} = 1.99$



Example: $s_1 = 19.24$, $s_2 = 4.35$, $E_{\mathrm{dim}} = 1.43$



Example: $s_1 = 83.65$, $s_2 = 4.33$, $E_{\text{dim}} = 1.10$



Conclusions

- Linear dynamical systems can be analyzed globally
- Nonlinear dynamical systems allow local analyses, but global ones are hard
- Weather (or at least a typical numerical weather model) is chaotic
- Ensembles are a Monte Carlo method for quantifying forecast uncertainty
- Local low dimensionality of chaotic models can be exploited for data assimilation