Generalized Local Linear Approximation

Steven M. Boker

Department of Psychology
University of Virginia

Dynamical Systems Analysis Workshop Part 3

Modern Modeling Methods
May 22, 2017

Overview

- Manifest variable two–step approaches to modeling derivatives.
- Local Linear Approximation (LLA) (Boker & Graham, 1998; Boker & Nesselroade, 2002; Boker, 2001).
- Generalized Local Linear Approximation (GLLA) (Boker, Deboeck, Edler, & Keel, 2010).
- Advantages and disadvantages of manifest variable methods.

A Linear First Order Differential Equation

- The simplest linear zero order system is

$$\dot{x}(t) = b_1$$

  where $\dot{x}(t)$ is the first derivative of the system at time t.
- The slope is a constant, b_1.
- The intercept isn't specified. So you need to pick an intercept to create a specific trajectory.
- To anthropomorphize:
  - No matter what the value of x or t, the slope of x is b_1.

The Specific Integral

- The specific integral of the previous zero order system is

$$x(t) = b_0 + b_1 t \ .$$

  where $\dot{x}(t)$ is the first derivative of the system at time t.
- The slope is a constant, b_1.
- Now we have an intercept because this is a specific trajectory.
- To anthropomorphize:
  - No matter what the value of x or t, the slope of x is b_1.
  - This system has an initial condition $(x(0))$ equal to b_0.

Another Linear First Order Differential Equation

- The simplest linear first order system is

$$\dot{x}(t) = b_1 x(t) \; .$$

  where $\dot{x}(t)$ is the first derivative of the system at time t.
- The slope $\dot{x}(t)$, is proportional to the value of x at time t.
- Again, the intercept doesn't matter.
- To anthropomorphize:
  - The farther x is from its equilibrium, the faster x changes.
  - If b_1 is negative, this is the equation of exponential decay.
  - Suppose x is a car and b_1 is negative, the farther the car is from its garage, the faster it goes back towards the garage.

The Specific Integral

- The specific integral of the previous linear first order system is

$$x(t) = b_0 + e^{b_1 t}$$

where $\dot{x}(t)$ is the first derivative of the system at time t.
- Now x is an exponential function of time.
- We have an asymptote because this is a specific trajectory.
- When $b_1 < 0$
  - The farther x is from its equilibrium, the faster x changes.
  - The asymptotic equilibrium is $\lim_{t \to \infty} x(t) = b_0$.

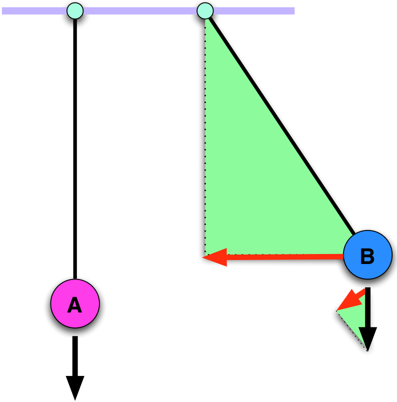
Linear Second Order Differential Equation

- A simple linear second order system is

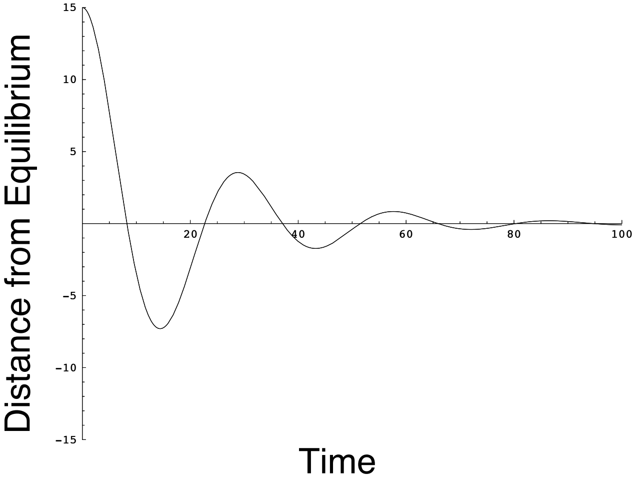
$$\ddot{x}(t) = \eta x(t) + \zeta \dot{x}(t)$$

  where x is the displacement from an equilibrium.
- If $\zeta < 0$, negative exponential damping.
- If $\eta < 0$ and $\eta + \zeta^2/4 < 0$, oscillation of period $\lambda = \frac{2\pi}{\sqrt{-(\eta+\zeta^2/4)}}$.
- To anthropomorphize, when $\zeta < 0$ and $\eta < 0$:
  - The farther x is from equilibrium, the more it wants to curve back towards equilibrium.
  - The faster x is changing, the more it wants to slow down.

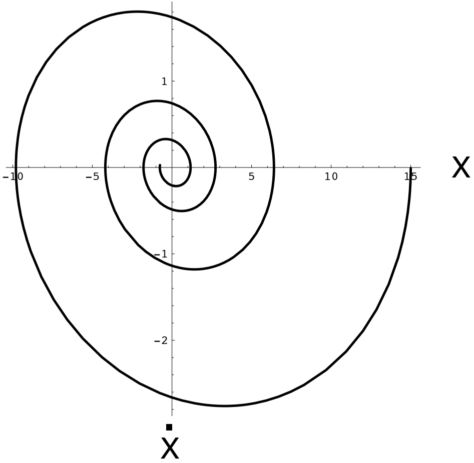
Passive Self–Regulation of a Pendulum



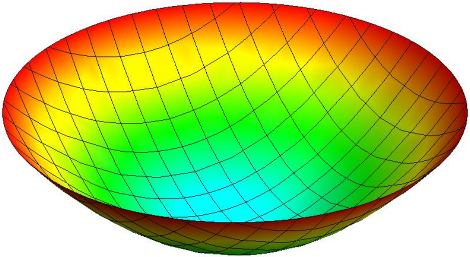
Acceleration due to gravity is proportional to displacement from equilibrium



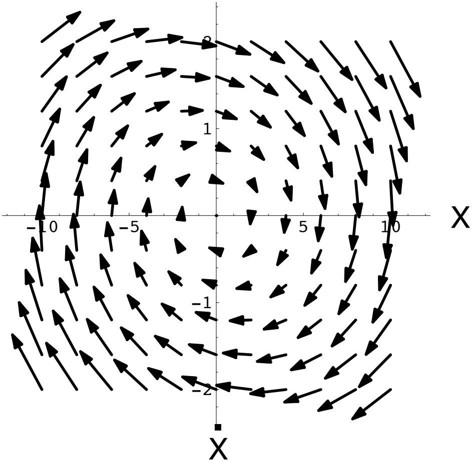
A Pendulum with Friction in a Gravitional Field



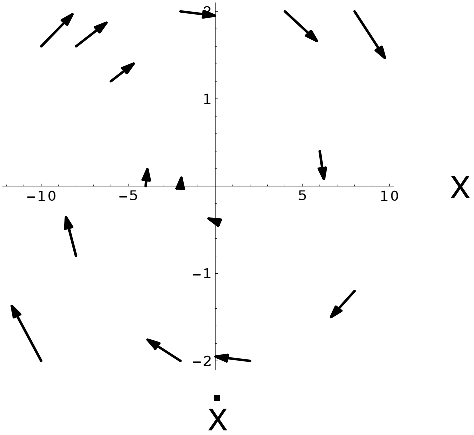
Displacement from equilibrium versus the first derivative of displacement



Attractor surface for the second order linear system



A vector field of the attractor of the second order linear system



A sample of ordered bursts from the vector field can be used to estimate its parameters

Estimating Differential Equations from Data

1. The data should be centered about the equilibrium
   - If you know the theoretical equilibrium, center about it.
   - If not, one popular method is to subtract the within–individual mean and linear trend.
   - Later in the workshop we will learn how to estimate individual-level equilibria simultaneously with the parameters of the differential equation.
2. Derivatives and displacements are estimated.
3. Regressions between derivatives and displacements are estimated.
4. Second level relationships are estimated.

Ideally, all of this happens simultaneously in one analysis procedure.

Two Convolution Filter Methods for Fitting Models

1. Generalized Local Linear Approximation (GLLA)
   - Can change the number of samples in the filter.
   - Reasonable rejection of measurement error.
   - For second order oscillating systems, breaks down if the measurement interval $> 1/4$ period
2. Latent Differential Equations (LDE)
   - Can use multivariate factor data.
   - Gives separate estimates of dynamic and measurement error.
   - For second order oscillating systems, breaks down if the measurement interval $> 1/4$ period, but a constrained 3rd order system can (sometimes) work with intervals up to $1/2$ the period.

Multi-Person Time Delay Embedding

- Suppose a time series X has been centered around each individual's equilibrium values.
- If the original time series X is ordered by occasion j within individual i then the series of all observations $x_{(i,j)}$ can be written as a vector of scores

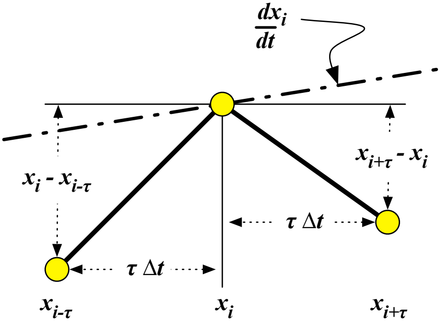
$$
\begin{aligned}
X \quad = \quad & \{x_{(1,1)}, x_{(1,2)}, \ldots x_{(1,P)}, x_{(2,1)}, x_{(2,2)}, \ldots x_{(2,P)}, \\
& \ldots, x_{(N,1)}, x_{(N,2)}, \ldots x_{(N,P)}\}.
\end{aligned}
$$

Time Delay Embedding

- For N people, each of whom have been sampled P times, a 5–D time delay embedded matrix $\mathbf{X}^{(5)}$ can be constructed as

$$\mathbf{X}^{(5)} = \begin{bmatrix} x_{(1,1)} & x_{(1,2)} & x_{(1,3)} & x_{(1,4)} & x_{(1,5)} \\ x_{(1,2)} & x_{(1,3)} & x_{(1,4)} & x_{(1,5)} & x_{(1,6)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{(1,P-4)} & x_{(1,P-3)} & x_{(1,P-2)} & x_{(1,P-1)} & x_{(1,P)} \\ x_{(2,1)} & x_{(2,2)} & x_{(2,3)} & x_{(2,4)} & x_{(2,5)} \\ x_{(2,2)} & x_{(2,3)} & x_{(2,4)} & x_{(2,5)} & x_{(2,6)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{(2,P-4)} & x_{(2,P-3)} & x_{(2,P-2)} & x_{(2,P-1)} & x_{(2,P)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{(N,1)} & x_{(N,2)} & x_{(N,3)} & x_{(N,4)} & x_{(N,5)} \\ x_{(N,2)} & x_{(N,3)} & x_{(N,4)} & x_{(N,5)} & x_{(N,6)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{(N,P-4)} & x_{(N,P-3)} & x_{(N,P-2)} & x_{(N,P-1)} & x_{(N,P)} \end{bmatrix}.$$

Local Linear Approximation



Linear Approximation with 3 Occasions of Measurement

Local Linear Approximation

1. Center each individual's data about his or her equilibrium.
2. For 3 observations $(x_{i-\tau}, x_i, x_{i+\tau})$ separated by a lag τ and interval between measurements Δt, the derivatives at x_i are

$$\begin{aligned}\dot{x}_i &= (x_{i+\tau} - x_{i-\tau})/2(\tau\Delta t)\\ \ddot{x}_i &= (x_{i+\tau} + x_{i-\tau} - 2x_i)/(\tau\Delta t)^2\end{aligned}$$

3. Then, use multiple regression

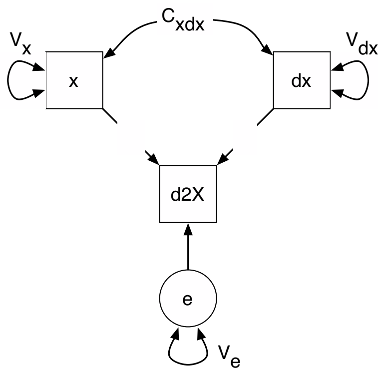
$$\ddot{x}_i = \eta x_i + \zeta \dot{x}_i + e_i$$

or a mixed effects multiple regression

$$\ddot{x}_{ij} = \eta_j x_{ij} + \zeta_j \dot{x}_{ij} + e_{ij}$$

or the derivatives can be used in a structural equation model.

Local Linear Approximation

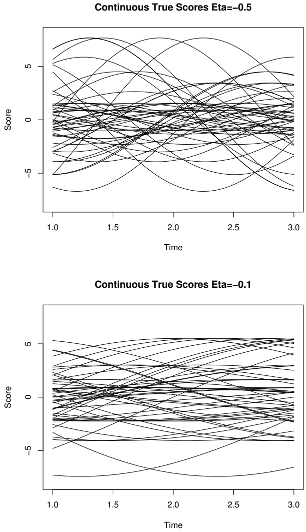


LLA estimates used as manifest variables in an SEM model.

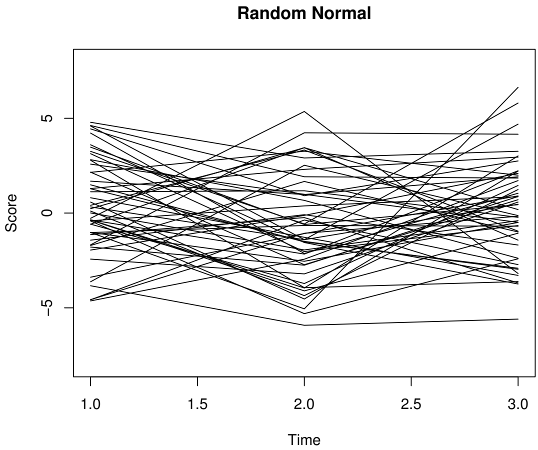
The Good, The Bad & The Ugly of LLA

- Good: LLA does a surprisingly good job of recovering true parameters in simulated data when analysis interval is optimal.
- Bad: When analysis interval isn't optimal, parameter estimate of η will be biased.
- Ugly: Normally distributed random numbers can be fit with this model, with an approximate expected $R^2 = 0.64$.

Three Occasions from Linear Oscillators



Three Independent Occasions from Normal Distribution



Correlations Between Derivatives for Noise

- 1000 normally distributed random numbers.
- Correlation between derivatives calculated with LLA.

x	$\dot{x}$	$\ddot{x}$
1.00	-0.04	-0.81
-0.04	1.00	0.06
-0.81	0.06	1.00

Why are x and $\ddot{x}$ related?

- Consider x_i drawn from normally distributed random numbers with mean of zero.
- No matter what the value of x_i, the expected values of x_{i-1} and x_{i+1} are zero.
- Thus the bigger x_i, the bigger $\ddot{x}_i$.

Suppose a null hypothesis of normally distributed zero mean random numbers was true. For any chosen occasion of measurement i we have a measured value x_i. Then,

$$\ddot{x}_i = \eta x_i + \zeta \dot{x}_i \tag{1}$$

$$\frac{x_{i+\tau} + x_{i-\tau} - 2x_i}{\tau^2} = \eta x_i + \zeta \frac{x_{i+\tau} - x_{i-\tau}}{\tau} \tag{2}$$

But

$$\mathcal{E}(x_{i+\tau}|x_i) = \mathcal{E}(x_{i-\tau}|x_i) = 0$$

so

$$\frac{0 + 0 - 2x_i}{\tau^2} = \eta x_i + \zeta \frac{0 - 0}{\tau} \tag{3}$$

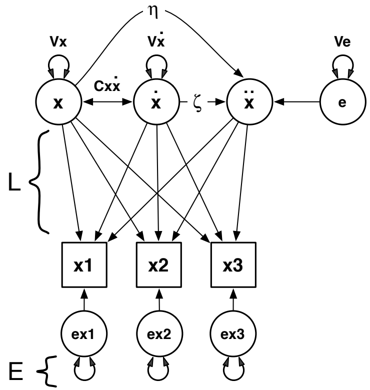
$$x_i \frac{-2}{\tau^2} = \eta x_i \tag{4}$$

$$\eta \tau^2 = -2 \tag{5}$$

Generalized Local Linear Approximation

- In order to address this problem, GLLA was developed.
- The basic idea is that a filter kernel wider than 3 will help reject time–independent noise.
- The fourth (or more) column in the embedding gives us a constraint that can smooth over time-independent noise.
- How should we build such a filter?
- In order to understand this, we first take a brief detour to Latent Differential Equations (LDE), a variation on Latent Growth Curves.
- We will come back to cover LDE models in detail in next week's lecture.

Brief Detour: Latent Differential Equations



Three lag embedded second order LDE Model (unidentified)

LDE Model Matrices

$$\mathbf{L} = \begin{bmatrix} 1 & -1\tau\Delta t & (-1\tau\Delta t)^2/2 \\ 1 & 0 & 0 \\ 1 & 1\tau\Delta t & (1\tau\Delta t)^2/2 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \eta & \zeta & 0 \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} V_F & C_{FdF} & 0 \\ C_{FdF} & V_{dF} & 0 \\ 0 & 0 & V_{d2F} \end{bmatrix}$$

$$\mathbf{R} = \mathbf{L}(\mathbf{I}-\mathbf{A})^{-1}\mathbf{S}(\mathbf{I}-\mathbf{A})^{-1'}\mathbf{L}' + \mathbf{U}^2$$

Three Dimensional Embedded Model

- Consider an LDE model fit to a time–delay embedded matrix $\mathbf{X}$.
- The general linear model that relates the latent differential scores $\mathbf{Y}$ to $\mathbf{X}$ can be written as

$$\mathbf{X} = \mathbf{Y}\mathbf{L}' + \mathbf{E}$$

- If $\mathbf{E} = 0$ and $\mathbf{L}'\mathbf{L}$ is nonsingular, we can solve for $\mathbf{Y}$

$$\begin{aligned} \mathbf{X}\mathbf{L} &= \mathbf{Y}\mathbf{L}'\mathbf{L} \\ \mathbf{X}\mathbf{L}(\mathbf{L}'\mathbf{L})^{-1} &= \mathbf{Y}(\mathbf{L}'\mathbf{L})(\mathbf{L}'\mathbf{L})^{-1} \\ \mathbf{X}\mathbf{L}(\mathbf{L}'\mathbf{L})^{-1} &= \mathbf{Y}\ . \end{aligned}$$

- So, if we define a matrix $\mathbf{W} = \mathbf{L}(\mathbf{L}'\mathbf{L})^{-1}$, then we can obtain derivative estimates with a single matrix multiplication

$$\mathbf{Y} = \mathbf{X}\mathbf{W}$$

The relationship between LDE and LLA

- Substituting the values from the matrix $\mathbf{L}$ defined in the previous slide, we can calculate a matrix $\mathbf{W}$ as

$$\begin{aligned} \mathbf{W} &= \mathbf{L}(\mathbf{L}'\mathbf{L})^{-1} \\ &= \begin{bmatrix} 0 & -1/2\tau\Delta t & 1/(\tau\Delta t)^2 \\ 1 & 0 & -2/(\tau\Delta t)^2 \\ 0 & 1/2\tau\Delta t & 1/(\tau\Delta t)^2 \end{bmatrix}. \end{aligned}$$

- The matrix $\mathbf{W}$ contains the coefficients from LLA such that the derivatives can be estimated as

$$\mathbf{Y} = \mathbf{X}\mathbf{W}\,.$$

Generalized Local Linear Approximation

- But now a transformation matrix $\mathbf{W}$ can be calculated for a time–delay embedding with any number columns, τ, and Δt.
- A 5–dimensional embedding $\mathbf{X}^{(5)}$ has a loading matrix

$$\mathbf{L} = \begin{bmatrix} 1 & -2\tau\Delta t & (-2\tau\Delta t)^2/2 \\ 1 & -1\tau\Delta t & (-1\tau\Delta t)^2/2 \\ 1 & 0 & 0 \\ 1 & 1\tau\Delta t & (1\tau\Delta t)^2/2 \\ 1 & 2\tau\Delta t & (2\tau\Delta t)^2/2 \end{bmatrix}.$$

If we choose $\tau = 1$ and $\Delta t = 1$,

$$\mathbf{W} = \begin{bmatrix} -0.0857 & -0.2 & 0.2857 \\ 0.3429 & -0.1 & -0.1429 \\ 0.4857 & 0.0 & -0.2857 \\ 0.3429 & 0.1 & -0.1429 \\ -0.0857 & 0.2 & 0.2857 \end{bmatrix}.$$

R Code to calculate the W matrix for GLLA

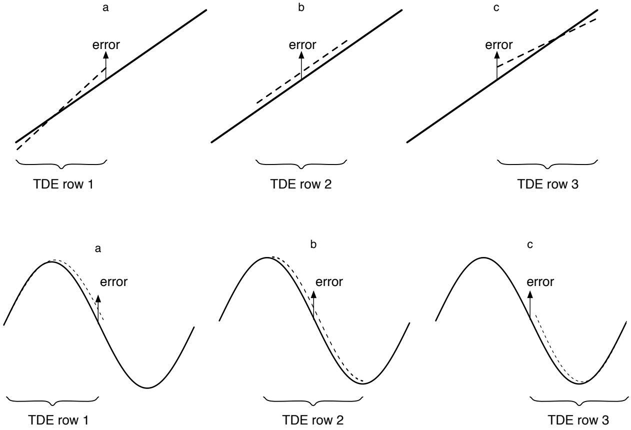
```
#-----------------------------------------------------------
# gllaWMatrix -- Calculates a GLLA linear transformation matrix to
#                create approximate derivatives
#
# Input:  embed -- Embedding dimension
#           tau -- Time delay
#        deltaT -- Interobservation interval
#         order -- Highest order of derivatives (2, 3, or 4)

gllaWMatrix <- function(embed, tau, deltaT, order=2) {
    L <- rep(1,embed)
    for(i in 1:order) {
        L <- cbind(L,(((c(1:embed)-mean(1:embed))*tau*deltaT)^i)/factorial(i))
    }
    return(L%*%solve(t(L)%*%L))
}
```

Advantages/Disadvantages of Time Delay Embedding

- Advantages:
  - Easy to use.
  - Surprisingly robust to non-equal sampling interval.
  - Empirical standard errors are smaller than those estimated by standard statistical software (Oertzen & Boker, 2010).
- Disadvantages.
  - Number of observations of dynamics decreases as number of columns in embedding matrix increases.
  - Empirical standard errors are smaller than those estimated by standard statistical software.

The Unreasonable Effectiveness of Time Delay Embedding



Advantages of GLLA

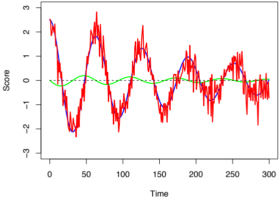
- Damping is recovered quite well.
- Frequency parameter is recovered well if there are sufficient observations in the kernel.
- No nonlinear constraints are required so models almost always converge.
- Off–the–shelf statistical techniques can be used, such as mixed effects modeling.

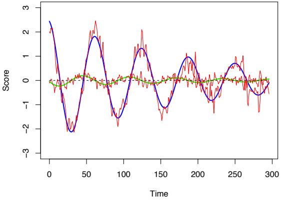
Problems with GLLA

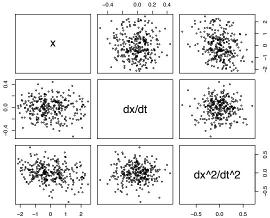
- Parameter values may be biased when sampling interval is not optimal.
- Noise can masquerade as fast frequency signal.
- No measurement model, thus measurement error cannot be distinguished from dynamic error.
- Not multivariate.

GLLA_SimExample160926.R

5 Dimensional Embedding

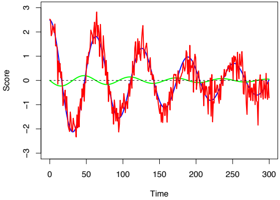


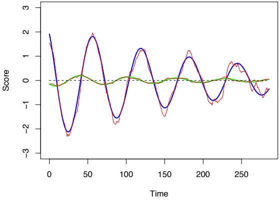


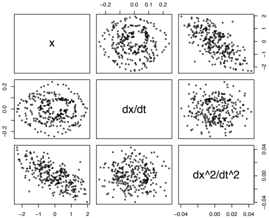


GLLA_SimExample160926.R

15 Dimensional Embedding

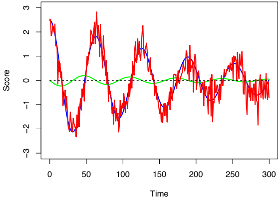


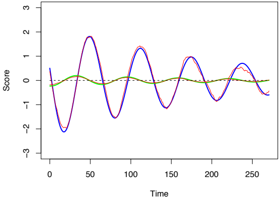


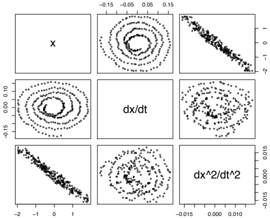


GLLA_SimExample160926.R

30 Dimensional Embedding







GLLA in Practice

- GLLA can recover parameters in simulation, and have been useful when applied to real world data.
- Examples include substance use in adolescents, rapid cycling bipolar disorder, and development of postural control in infants.
- A further example is data from a study of the grieving process in recent widows.

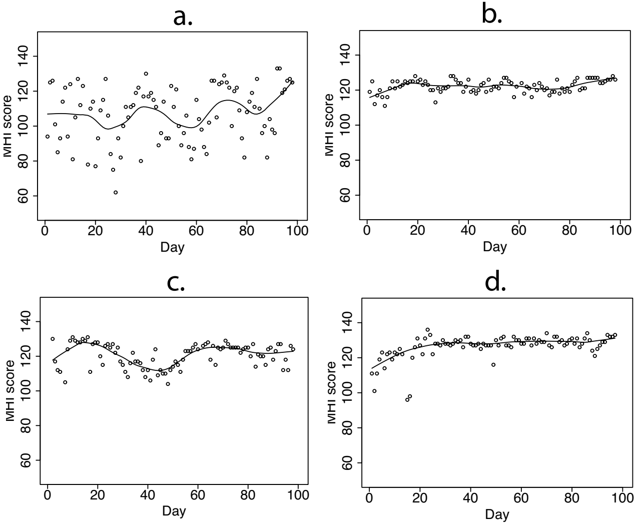
Self–Reported Mental Health in Recent Widows

- Recent widows reported a Mental Health Inventory (**MHI**) daily for 90 days.
- Other variables were collected in pre- and post-interviews, including three pretest variables: Perceived Control, Emotion Focused Social Support, and Problem Focused Social Support.

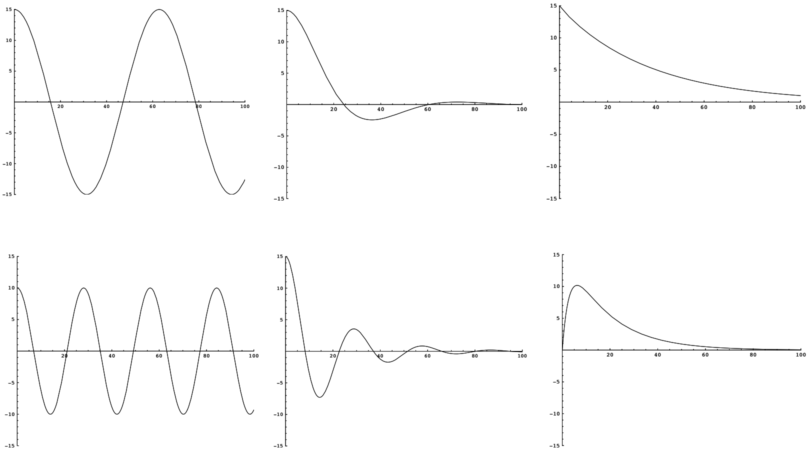
Mental Health Inventory

- Mental Health Inventory (MHI) was selected in part due to its relatively high internal consistency (in the range .83 to .91) but relatively low test-retest correlation (in the range .56 to .64).
- Considerable intraindividual variability was observed in self reported MHI scores over the course of 90 days.

Example MHI Scores with Loess Smooth



Example Linear Oscillator Trajectories



Interindividual Differences in Intraindividual Trajectories

- A random coefficients model.

$$
\begin{aligned}
\ddot{x}_{ij} &= \eta_i x_{ij} + \zeta_i \dot{x}_{ij} + e_{ij} \\
\eta_i &= c_{00} + u_{0i} \\
\zeta_i &= c_{10} + u_{1i}
\end{aligned}
$$

where x_{ij} is the ith person's MHI score at the jth occasion.

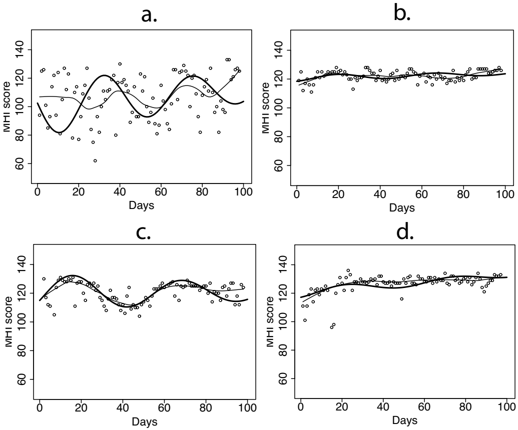
Results of Multilevel Model ($\tau = 12$)

AIC	BIC	logLik
-3423.362	-3391.816	1717.681

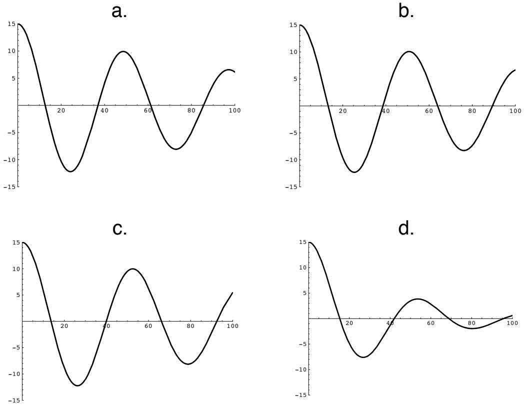
	Value	SE	t	p
η	-0.0149	0.0004	-36.27	$< .001$
ζ	-0.0198	0.0057	-3.48	$< .001$

Number of Observations: 1421
Number of Groups: 28

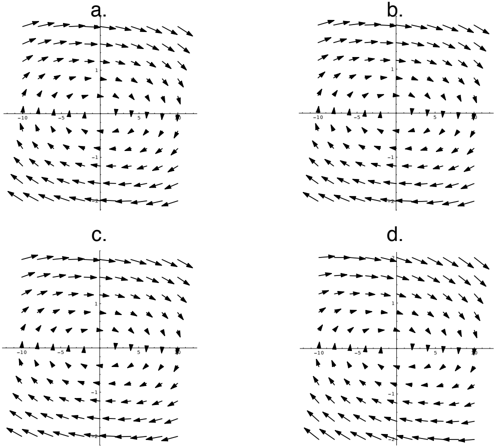
Example MHI Scores with Model Trajectories



Trajectories with Equal Initial Conditions



Vector Field Plots



Interindividual Differences in Intraindividual Trajectories

- Second level random coefficients model.

$$
\begin{aligned}
\ddot{x}_{ij} &= \eta_i x_{ij} + \zeta_i \dot{x}_{ij} + e_{ij} \\
\eta_i &= c_{00} + c_{01} y_i + c_{02} z_i + c_{03} w_i + u_{0i} \\
\zeta_i &= c_{10} + c_{11} y_i + c_{12} z_i + c_{13} w_i + u_{1i}
\end{aligned}
$$

where x_{ij} is the ith persons MHI score at the jth occasion, y_i is Perceived Control, z_i is Problem Focused Coping and w_i is Emotion Focused Coping provided by the family of person i.

Results of Multilevel Model ($\tau = 11$)

AIC	BIC	logLik	DF
-2576	-2514	-2600	1302

	Value	SE	t	p
ζ	-0.02399	0.00502	-4.779	$< .0001$
$PC \rightarrow \zeta$	-0.00142	0.00124	-1.145	0.2524
$PFC \rightarrow \zeta$	-0.00100	0.00178	-0.561	0.5742
$EFC \rightarrow \zeta$	-0.00267	0.00128	-2.091	0.0367
η	-0.01659	0.00042	-39.270	$< .0001$
$PC \rightarrow \eta$	0.00026	0.00010	2.616	0.0090
$PFC \rightarrow \eta$	-0.00018	0.00016	-1.128	0.2592
$EFC \rightarrow \eta$	0.00014	0.00013	1.102	0.2707

Conclusions from Widows' Data

- The residual variability in the MHI could be modeled as a damped linear oscillator.
- There was evidence of interindividual differences in the parameters of this model.
- Higher values of Emotion Focused coping were associated with quicker damping of oscillations.
- Higher values of Perceived Control were associated with slightly slower oscillations.

Thank You

Boker, S. M. (2001). Differential structural modeling of intraindividual variability. In L. Collins & A. Sayer (Eds.), *New methods for the analysis of change* (pp. 3–28). Washington, DC: APA.

Boker, S. M., Deboeck, P. R., Edler, C., & Keel, P. K. (2010). Generalized local linear approximation of derivatives from time series. In S.-M. Chow & E. Ferrar (Eds.), *Statistical methods for modeling human dynamics: An interdisciplinary dialogue* (pp. 161–178). Boca Raton, FL: Taylor & Francis.

Boker, S. M., & Graham, J. (1998). A dynamical systems analysis of adolescent substance abuse. *Multivariate Behavioral Research*, *33*(4), 479–507.

Boker, S. M., & Nesselroade, J. R. (2002). A method for modeling the intrinsic dynamics of intraindividual variability: Recovering the parameters of simulated oscillators in multi–wave panel data. *Multivariate Behavioral Research*, *37*(1), 127–160.

Oertzen, T. v., & Boker, S. M. (2010). Time delay embedding increases estimation precision of models of intraindividual variability. *Psychometrika*, *75*(1), 158–175.