Introduction to Nonlinear Time Series Analysis

“Using a term like nonlinear science is like referring to the bulk of zoology as the study of non-elephant animals.”
Nonlinear Dynamics in the Classroom
Nonlinear Dynamics in the Classroom

\[
\frac{dx}{dt} = \sigma(y - x), \\
\frac{dy}{dt} = x(\rho - z) - y, \\
\frac{dz}{dt} = xy - \beta z.
\]

\[
x_{n+1} = 1 - ax_n^2 + y_n \\
y_{n+1} = bx_n.
\]
Nonlinear Dynamics in the Wild

\[
\begin{align*}
\frac{dx}{dt} &= \sigma(y - x), \\
\frac{dy}{dt} &= r(x - p) - y, \\
\frac{dz}{dt} &= xy - \beta z.
\end{align*}
\]

\[
\begin{align*}
x_{n+1} &= 1 - ax_n^2 + y_n \\
y_{n+1} &= bx_n.
\end{align*}
\]
Observation

“Observe the system”
Or
“Collect a Time Series”
Observation

Temperature each day
Number of phytoplankton at time $n$
Number of infected in population
The Challenge
The Goal

“Inflate” observations of the system into equivalent system
Delay Coordinate Embedding

OBSERVE

RECONSTRUCT
Delay Coordinate Embedding Motivation

Been used to successfully **explore, predict and understand** many diverse complex systems

- Roulette Wheels (The prediction company)
- Traded Financial Markets (The prediction company)
- Phytoplankton Populations
- Computer Performance Dynamics (JG & E. Bradley)
- SFI A (Far-Infrared-Laser)
- Disease Outbreaks
- Ground Water Levels
- …
Aside … Topological Data Analysis…
Delay Coordinate Embedding
### Delay Coordinate Embedding

How can we take these observations and faithfully reconstruct the underlying dynamics?!

<table>
<thead>
<tr>
<th>$\hat{x}$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-78.531</td>
<td>0.00</td>
</tr>
<tr>
<td>-74.110</td>
<td>0.01</td>
</tr>
<tr>
<td>-70.385</td>
<td>0.02</td>
</tr>
<tr>
<td>-67.286</td>
<td>0.03</td>
</tr>
<tr>
<td>-64.760</td>
<td>0.04</td>
</tr>
<tr>
<td>-62.756</td>
<td>0.05</td>
</tr>
<tr>
<td>-61.218</td>
<td>0.06</td>
</tr>
<tr>
<td>-60.104</td>
<td>0.07</td>
</tr>
<tr>
<td>-59.389</td>
<td>0.08</td>
</tr>
<tr>
<td>-59.036</td>
<td>0.09</td>
</tr>
<tr>
<td>-59.032</td>
<td>0.10</td>
</tr>
<tr>
<td>-59.357</td>
<td>0.11</td>
</tr>
<tr>
<td>-60.009</td>
<td>0.12</td>
</tr>
<tr>
<td>-60.983</td>
<td>0.13</td>
</tr>
<tr>
<td>-62.273</td>
<td>0.14</td>
</tr>
<tr>
<td>-63.920</td>
<td>0.15</td>
</tr>
<tr>
<td>-65.903</td>
<td>0.16</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>
Delay Coordinate Embedding

\[ \hat{x}(t) \]

\[ \hat{x}(t - \tau) \]

\[ \hat{x}(t - 2\tau) \]

\[ (\hat{x}(t - \tau), \hat{x}(t), \hat{x}(t + \tau)) \]

\[ \tau = 0.05 \]

\[ m = 3 \]

\[ (-59.032, -62.756, -78.531) \]

\( \tau \) is called the “time delay”

\( m \) is called the “embedding dimension”

\( \hat{x} \)

\( t \)

\( \bar{x} \)
Delay Coordinate Embedding

<table>
<thead>
<tr>
<th>$\hat{x}$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-78.531</td>
<td>0.00</td>
</tr>
<tr>
<td>-74.110</td>
<td>0.01</td>
</tr>
<tr>
<td>-70.385</td>
<td>0.02</td>
</tr>
<tr>
<td>-67.286</td>
<td>0.03</td>
</tr>
<tr>
<td>-64.760</td>
<td>0.04</td>
</tr>
<tr>
<td>-62.756</td>
<td>0.05</td>
</tr>
<tr>
<td>-61.218</td>
<td>0.06</td>
</tr>
<tr>
<td>-60.104</td>
<td>0.07</td>
</tr>
<tr>
<td>-59.389</td>
<td>0.08</td>
</tr>
<tr>
<td>-59.036</td>
<td>0.09</td>
</tr>
<tr>
<td>-59.032</td>
<td>0.10</td>
</tr>
<tr>
<td>-59.357</td>
<td>0.11</td>
</tr>
<tr>
<td>-60.009</td>
<td>0.12</td>
</tr>
<tr>
<td>-60.983</td>
<td>0.13</td>
</tr>
<tr>
<td>-62.273</td>
<td>0.14</td>
</tr>
<tr>
<td>-63.920</td>
<td>0.15</td>
</tr>
<tr>
<td>-65.903</td>
<td>0.16</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

$\hat{x}(t - \tau)$

$\hat{x}(t)$

$\hat{x}(t - 2\tau)$

$\tau = 0.05$

$m = 3$

$\tau$ is called the “time delay”

$m$ is called the “embedding dimension”
**Delay Coordinate Embedding**

<table>
<thead>
<tr>
<th>( \hat{x} )</th>
<th>( t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-78.531</td>
<td>0.00</td>
</tr>
<tr>
<td>-74.110</td>
<td>0.01</td>
</tr>
<tr>
<td>-70.385</td>
<td>0.02</td>
</tr>
<tr>
<td>-67.286</td>
<td>0.03</td>
</tr>
<tr>
<td>-64.760</td>
<td>0.04</td>
</tr>
<tr>
<td>-62.756</td>
<td>0.05</td>
</tr>
<tr>
<td>-61.218</td>
<td>0.06</td>
</tr>
<tr>
<td>-60.104</td>
<td>0.07</td>
</tr>
<tr>
<td>-59.389</td>
<td>0.08</td>
</tr>
<tr>
<td>-59.036</td>
<td>0.09</td>
</tr>
<tr>
<td>-59.032</td>
<td>0.10</td>
</tr>
<tr>
<td>-59.357</td>
<td>0.11</td>
</tr>
<tr>
<td>-60.009</td>
<td>0.12</td>
</tr>
<tr>
<td>-60.983</td>
<td>0.13</td>
</tr>
<tr>
<td>-62.273</td>
<td>0.14</td>
</tr>
<tr>
<td>-63.920</td>
<td>0.15</td>
</tr>
<tr>
<td>-65.903</td>
<td>0.16</td>
</tr>
</tbody>
</table>

\( \tau = 0.05 \)

\( m = 3 \)

\( \hat{x}(t) \)

\( \hat{x}(t - \tau) \)

\( \hat{x}(t - 2\tau) \)

\( \tau \) is called the “time delay”

\( m \) is called the “embedding dimension”
## Delay Coordinate Embedding

### Table:

<table>
<thead>
<tr>
<th>( \hat{x} )</th>
<th>( t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-78.531</td>
<td>0.00</td>
</tr>
<tr>
<td>-74.110</td>
<td>0.01</td>
</tr>
<tr>
<td>-70.385</td>
<td>0.02</td>
</tr>
<tr>
<td>-67.286</td>
<td>0.03</td>
</tr>
<tr>
<td>-64.760</td>
<td>0.04</td>
</tr>
<tr>
<td>-62.756</td>
<td>0.05</td>
</tr>
<tr>
<td>-61.218</td>
<td>0.06</td>
</tr>
<tr>
<td>-60.104</td>
<td>0.07</td>
</tr>
<tr>
<td>-59.389</td>
<td>0.08</td>
</tr>
<tr>
<td>-59.036</td>
<td>0.09</td>
</tr>
<tr>
<td>-59.032</td>
<td>0.10</td>
</tr>
<tr>
<td>-59.357</td>
<td>0.11</td>
</tr>
<tr>
<td>-60.009</td>
<td>0.12</td>
</tr>
<tr>
<td>-60.983</td>
<td>0.13</td>
</tr>
<tr>
<td>-62.273</td>
<td>0.14</td>
</tr>
<tr>
<td>-63.920</td>
<td>0.15</td>
</tr>
<tr>
<td>-65.903</td>
<td>0.16</td>
</tr>
</tbody>
</table>

### Diagram:

- \( \tau = 0.05 \)
- \( m = 3 \)

...continue for the entire time series
Takens* Theorem

For the right $\tau$ and enough dimensions*, the embedded dynamics are diffeomorphic to (have same topology as) the original state-space dynamics.

This idea was originally formulated by Norman Packard, J. Doyne Farmer, James Crutchfield, and R. Shaw but formalized later by *Whitney, Mane, …,

*and an infinitely long noise free time series
Diffeomorphisms and Topology

What that means:

• Qualitatively the same topological shape

• Same dynamical invariants (e.g., $\lambda$, fractal dimension)

The climate…

…perfectly!
Takens* Theorem

For the right $\tau$ and enough dimensions*, the embedded dynamics are diffeomorphic to (have same topology as) the original state-space dynamics.

This idea was originally formulated by Norman Packard, J. Doyne Farmer, James Crutchfield, and R. Shaw but formalized later by *Whitney, Mane, …,

*and an infinitely long noise free time series
The Observation Function

$h(\tilde{x})$ needs to be a smooth function of the unknown state space variables.

THOUGHT EXPERIMENT…What functions would and would not work?

\[
\begin{align*}
\dot{x} &= \sigma(y - x), \\
\dot{y} &= x(\rho - z) - y, \\
\dot{z} &= xy - \beta z.
\end{align*}
\]

<table>
<thead>
<tr>
<th>(\dot{x})</th>
<th>(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-78.531</td>
<td>0.00</td>
</tr>
<tr>
<td>-74.110</td>
<td>0.01</td>
</tr>
<tr>
<td>-70.385</td>
<td>0.02</td>
</tr>
<tr>
<td>-67.286</td>
<td>0.03</td>
</tr>
<tr>
<td>-64.760</td>
<td>0.04</td>
</tr>
<tr>
<td>-62.756</td>
<td>0.05</td>
</tr>
<tr>
<td>-61.218</td>
<td>0.06</td>
</tr>
<tr>
<td>-60.104</td>
<td>0.07</td>
</tr>
<tr>
<td>-59.389</td>
<td>0.08</td>
</tr>
<tr>
<td>-59.036</td>
<td>0.09</td>
</tr>
<tr>
<td>-59.032</td>
<td>0.10</td>
</tr>
<tr>
<td>-59.357</td>
<td>0.11</td>
</tr>
<tr>
<td>-60.009</td>
<td>0.12</td>
</tr>
<tr>
<td>-60.983</td>
<td>0.13</td>
</tr>
<tr>
<td>-62.273</td>
<td>0.14</td>
</tr>
<tr>
<td>-63.920</td>
<td>0.15</td>
</tr>
<tr>
<td>-65.903</td>
<td>0.16</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
</tr>
</tbody>
</table>

\[ F(\tau, m) \]

\[ x(t - \tau) \] 

\[ x(t) \]
The Observation Function

\( h(\vec{x}) \) needs to be a smooth function of the unknown state space variables.

\[ h \in C^1(\mathbb{R}^d \to \mathbb{R}) \quad \text{e.g.,} \quad \hat{x} = h([x \ y \ z]^T) = xy - z \]

or \( \hat{x} = h([x \ y \ z]^T) = x \)

\[ \dot{x} = \sigma(y - x), \]
\[ \dot{y} = x(\rho - z) - y, \]
\[ \dot{z} = xy - \beta z. \]

<table>
<thead>
<tr>
<th>( \hat{x} )</th>
<th>( t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-78.531</td>
<td>0.00</td>
</tr>
<tr>
<td>-74.110</td>
<td>0.01</td>
</tr>
<tr>
<td>-70.385</td>
<td>0.02</td>
</tr>
<tr>
<td>-67.286</td>
<td>0.03</td>
</tr>
<tr>
<td>-64.760</td>
<td>0.04</td>
</tr>
<tr>
<td>-62.756</td>
<td>0.05</td>
</tr>
<tr>
<td>-61.218</td>
<td>0.06</td>
</tr>
<tr>
<td>-60.104</td>
<td>0.07</td>
</tr>
<tr>
<td>-59.389</td>
<td>0.08</td>
</tr>
<tr>
<td>-59.036</td>
<td>0.09</td>
</tr>
<tr>
<td>-59.032</td>
<td>0.10</td>
</tr>
<tr>
<td>-59.357</td>
<td>0.11</td>
</tr>
<tr>
<td>-60.009</td>
<td>0.12</td>
</tr>
<tr>
<td>-60.983</td>
<td>0.13</td>
</tr>
<tr>
<td>-62.273</td>
<td>0.14</td>
</tr>
<tr>
<td>-63.920</td>
<td>0.15</td>
</tr>
<tr>
<td>-65.903</td>
<td>0.16</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
</tr>
</tbody>
</table>

THOUGHT EXPERIMENT…What function would and would not work..
The Time Delay (the “easy one”)

• How far apart the teeth in the “comb” are

• How far apart each coordinate is temporally (how far apart the axis of the reconstruction space are)

• Loose theoretical restriction on time delay, just needs to be:
  • positive (and not a multiple of the orbit’s period)

• In practice however, due to finite precision and arithmetic error very important

\[ \hat{x} = h(\bar{x}) \]

<table>
<thead>
<tr>
<th>( \hat{x} )</th>
<th>( t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-78.531</td>
<td>0.00</td>
</tr>
<tr>
<td>-74.110</td>
<td>0.01</td>
</tr>
<tr>
<td>-70.385</td>
<td>0.02</td>
</tr>
<tr>
<td>-67.286</td>
<td>0.03</td>
</tr>
<tr>
<td>-64.760</td>
<td>0.04</td>
</tr>
<tr>
<td>-62.756</td>
<td>0.05</td>
</tr>
<tr>
<td>-61.218</td>
<td>0.06</td>
</tr>
<tr>
<td>-60.104</td>
<td>0.07</td>
</tr>
<tr>
<td>-59.389</td>
<td>0.08</td>
</tr>
<tr>
<td>-59.036</td>
<td>0.09</td>
</tr>
<tr>
<td>-59.032</td>
<td>0.10</td>
</tr>
<tr>
<td>-59.357</td>
<td>0.11</td>
</tr>
<tr>
<td>-60.009</td>
<td>0.12</td>
</tr>
<tr>
<td>-60.983</td>
<td>0.13</td>
</tr>
<tr>
<td>-62.273</td>
<td>0.14</td>
</tr>
<tr>
<td>-63.920</td>
<td>0.15</td>
</tr>
<tr>
<td>-65.903</td>
<td>0.16</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
</tr>
</tbody>
</table>

\[ F(\tau, m) \]

\( \tau \) is called the “time delay”
Choosing Time Delay

• GOAL: Each delay coordinate tells you something new (but relevant)

• Redundancy versus Irrelevance (very delicate balance)
  • Want to minimize redundancy between variables
  • Want the variables to be causally related (not irrelevant)

• THOUGHT EXPERIMENT
  • Want to predict today’s temperature…
    • Do you want to know yesterday’s temperature? (tau = 1 day)
    • The Temperature a week ago? (tau = 1 week)
    • 74 years ago? (tau = 74 years)
Choosing Time Delay

• GOAL: Each delay coordinate tells you something new (but relevant)
• Redundancy versus Irrelevance (very delicate balance)
  • Want to minimize redundancy between variables
  • Want the variables to be causally related (not irrelevant)
• THOUGHT EXPERIMENT
  • Want to predict the next seconds closing price of Apple…
  • The last second’s price? (tau = 1 second)
  • Do you want to know yesterday’s price? (tau = 1 day)
  • The price a week ago? (tau = 1 week)
  • 74 years ago? (tau = 74 years)
Choosing Time Delay

• First approximation ---- Seek Independence from Dependence
  • Seek a time delay that results in independent coordinates

• THOUGHT EXPERIMENT
  • How can we accomplish this?
Choosing Time Delay

• First approximation ---- Seek Independence from Dependence
  • Seek a time delay that results in independent coordinates

  • How can we accomplish this?

\[ R(\tau) = \frac{1}{N - \tau} \frac{\sum_j (x_j - \mu_x)(x_{j-\tau} - \mu_x)}{\sigma^2_x} \]

• THOUGHT EXPERIMENT Why is this bad?
Choosing Time Delay

• First approximation ---- Seek Independence from Dependence
  • Seek a time delay that results in independent coordinates

• How can we accomplish this?

\[ R(\tau) = \frac{1}{N - \tau} \sum_j (x_j - \mu_x)(x_{j-\tau} - \mu_x) \frac{\sigma_x^2}{\sigma_x^2} \]

• THOUGHT EXPERIMENT Why is this bad?
Choosing Time Delay

• Second approximation —— Seek Independence and take into account nonlinear…

• Instead seek “General Independence” (Minimum mutual information)

\[ I[X_{j-\tau}; X_{j}] = \sum_{x_{j-\tau}, x_{j}} p(x_{j-\tau}, x_{j}) \log \frac{p(x_{j-\tau}, x_{j})}{p(x_{j-\tau})p(x_{j})} \]
Nonlinear time-series analysis revisited

Elizabeth Bradley\textsuperscript{1,\textsuperscript{a}) and Holger Kantz\textsuperscript{2,\textsuperscript{b})

\textsuperscript{1}Department of Computer Science, University of Colorado, Boulder, Colorado 80309-0430, USA and Santa Fe Institute, Santa Fe, New Mexico 87501, USA

\textsuperscript{2}Max Planck Institute for the Physics of Complex Systems, Noethnitzer Str. 38 D, 01187 Dresden, Germany

(Received 27 January 2015; accepted 26 March 2015; published online 13 April 2015)
Amplitude.dat

```
data
0 500 1000 1500 2000 2500 3000 3500 4000 4500 5000
-20
-10
0
10
20
```

![Graph of Amplitude.dat](image)
THOUGHT EXPERIMENT

So what should the delay be?
Why does it go back up?
Choosing Time Delay

$I(\text{bits})$ vs. $\tau(\times 0.001\,\text{s})$
Choosing Time Delay
We have found the delay!

Ok we have selected the “easy” parameter…
Any Questions?

```
mutual ./amplitude.dat -o
```
In theory need to choose $m > 2d_{cap}$ to obtain topological conjugacy (may occur sooner)

\[
\hat{x} = h(\tilde{x})
\]

$\hat{x}$ | $t$
--- | ---
-78.531 | 0.00
-74.110 | 0.01
-70.385 | 0.02
-67.286 | 0.03
-64.760 | 0.04
-62.756 | 0.05
-61.218 | 0.06
-60.104 | 0.07
-59.389 | 0.08
-59.036 | 0.09
-59.032 | 0.10
-59.357 | 0.11
-60.009 | 0.12
-60.983 | 0.13
-62.273 | 0.14
-63.920 | 0.15
-65.903 | 0.16

\[F(\tau, m)\]

$m$ is called the “embedding dimension”
Embedding Dimension

- In theory need to choose $m > 2d_{cap}$ to obtain topological conjugacy (may occur sooner)
  - Enough dimensions to stretch into, think about bridge
  - THOUGHT EXPERIMENT What is the logic behind this?

How can we tell these systems are being plotted in too few dimensions?
Embedding Dimension

- In theory need to choose $m > 2d_{\text{cap}}$ to obtain topological conjugacy (may occur sooner)
  - Enough dimensions to stretch into, think about bridge
  - THOUGHT EXPERIMENT What is the logic behind this?

$$p(\text{collision}) \propto \epsilon^{m-2d}$$

How can we tell these systems are being plotted in too few dimensions?
Embedding Dimension

- So we need $m > 2d_{cap}$
  - But we do not know the dimension …
  - BUT! Notion of enough dimensions to stretch into, think about the bridge
- Two standard approaches
  - Method of False neighbors
  - Method of Dynamical Invariants

$m$ is called the “embedding dimension”
Deterministic Dynamics can NOT cross in state space!

So simply figure out if dynamics cross and if they do this means embedding dimension is not large enough
False Neighbors
False Neighbors TISEAN

false_nearest -D8 -M1,8 ./amplitude.dat -o
False Neighbors TISEAN

```
false_nearest -D8 -M1,8 ./amplitude.dat -o
```
**False Neighbors TISEAN**

\[
\text{false\textunderscore nearest -D8 -M1,8 ./amplitude.dat -o}
\]

Embedding Dimension \( m \)

Percent of false neighbors

\( 3 \leq m \leq 6 \)
Delay Coordinate Embedding

delay -d8 -m3 ./amplitude.dat -o

![Amplitude vs. Time](chart1)

![Delay Coordinate Representation](chart2)
Delay Coordinate Embedding

delay -d8 -m3 ./amplitude.dat -o
Delay Coordinate Embedding

We have reconstructed the dynamics!

Selecting the time delay how?

Selecting the dimension how?

...ok so what? Just pretty pictures?
Delay Coordinate Embedding is useful!

Been used to successfully explore, predict and understand many diverse complex systems

- Roulette Wheels (The prediction company)
- Traded Financial Markets (The prediction company)
- Phytoplankton Populations
- Computer Performance Dynamics (JG & E. Bradley)
- SFI A (Far-Infrared-Laser)
- Disease Outbreaks
- Ground Water Levels
- …
Are the markets fractal?

Is a bear or bull market more space filling?

What is the topology of corruption?
Fractal Dimension

A point is zero dimensional
A line is one dimensional
A plane is two dimensional
A Cantor set is ??? dimensional
Correlation Dimension

How much of space is taken up by an object?

The correlation sum..

\[ C(\epsilon; m, \tau) = \frac{1}{N^2} \sum_{i,j=1, i \neq j}^{N} \Theta[\epsilon - ||\vec{x}_i - \vec{x}_j||] \]

\[ \Theta(x) = \begin{cases} 
1 & : x > 0 \\
0 & : x \leq 0 
\end{cases} \]

\[ C(\epsilon; m, \tau) \propto \epsilon^D \]

Slope of scaling region on this log-log plot is the Correlation Dimension
Correlation Dimension

\[ C(\varepsilon; m, \tau) \]

\[ \varepsilon \]
Correlation Dimension

\[ D \approx 2.1 \]

\[ 4 \leq m \leq 6 \]
Correlation Dimension

\[ D \approx 2.1 \]

\[ 4 \leq m \leq 6 \]
Lyapunov Exponent

```
lyap_k amplitude.dat -d8 -M3,6 -t100 -r.1 -s500 -o
```
Lyapunov Exponent

$$\lambda_1 = 0.015 \pm 0.002$$

$$4 \leq m \leq 6$$
WARNING!

Image taken from “Computer Systems are Dynamical Systems”
Delay Coordinate Embedding Takeaways

Nonlinear Time Series Analysis is **Hard and Subjective** … but!

"Using a term like **nonlinear** science is like referring to the bulk of zoology as the study of **non-elephant** animals."

- In "non-elephant" systems nonlinear tools are needed!
- You now know how to use the workhorse of this field.
Thank you and any questions?