University of Colorado
Department of Computer Science

Chaotic Dynamics – CSCI 4446/5446
Spring 2012

Problem Set 8

Issued:	6 March 2012
Due:	13 March 2012

Graduate Students: project paragraphs are due in class on 15 March!

Reading: *Sections 12.4–12.5 of Strogatz, section 3 of Liz's TSA Notes, and sections 3.2-3.3 and 9.1-9.2 of Kantz & Schreiber*. Other useful references for delay-coordinate embedding include chapters 2, 3, and 4 of the Abarbanel text and chapters 5 and 6 of "Coping with Chaos," both of which are on library reserve.

Bibliography:

- H. Abarbanel, *Analysis of Observed Chaotic Data*, Springer, 1995. A great general book on time-series analysis, from embedding onwards. This is the optional text for the class, and is on library reserve.
- A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical Review A* **33**:1134-1140 (1986). Average mutual information.
- C. Hsu, *Cell-to-Cell Mapping*, Springer-Verlag, 1987.
- J. P. Huke, "Embedding Nonlinear Dynamical Systems: A Guide to Takens' Theorem," Manchester preprint, 2006. See the link on the course webpage.
- M. B. Kennel *et al.*, "Determining minimum embedding dimension using a geometrical construction," *Physical Review A*, **45**:3403-3411 (1992). False near neighbors; this paper is in the "Coping..." collection. (A synopsis of this algorithm appears on page 17 of Liz's TSA Notes.)
- G. B. Mindlin and R. Gilmore, "Topological Analysis of Chaotic Time Series Data" in *Proceedings of the First Experimental Chaos Conference*, World Scientific, 1991. Like FNN, but based on topology instead of geometry.
- N. Packard *et al.*, "Geometry from a Time Series," *Physical Review Letters* **45**:712 (1980). One of the two original papers on embedding; also in the "Coping..." collection.
- T. Sauer *et al.*, "Embedology," *Journal of Statistical Physics*, **65**:579-616 (1991). The most useful paper about the broader field of embedding.

- T. Sauer, "Interspike interval embedding of chaotic signals," *Chaos*, **5**:127 (1995). An embedding method that uses the intervals between discrete events, rather than evenly-sampled measurements of the whole time line.

- F. Takens, "Detecting strange attractors in fluid turbulence," in *Dynamical Systems and Turbulence* (D. Rand and L.-S. Young, eds.), Springer, Berlin, 1981. The other original paper on reconstruction.

Problems:

In this problem set, you will explore some simple embedding algorithms, using position-versus-time data gathered from a real driven pendulum. I have posted three data sets on the class webpage; see the PS8 entry on that page for directions (and a clickable link) to these data. In all three runs, the angle was measured every Δt seconds using an optical encoder with a resolution of 0.4 degree. The drive amplitude was fixed; the drive frequency (the bifurcation parameter) was different for each data set:

- in `data1`, the drive was turned off
- in `data2`, the drive is on, with a medium frequency
- in `data3`, the drive is on, with the same amplitude but a higher frequency

> *Size Issues:* These data files contain up to 6MB of information. It would make sense to debug your code on test files that consist of *small chunks* of these large files.

Each file captures a single trajectory of the driven pendulum (except that data2 was so big that I broke it down into four pieces: `data2.first250sec`, `data2.second250sec`, and so on). Each line of each file represents a single time-sample of the pendulum's angular position. Each of these data points looks like this:

θ `time`

...where time is in seconds and θ is mod 2π. Depending on when I hit the reset button, θ may contain an offset, so "θ=0" may not be "vertical." Also, note that the sampling *rate* was different; `data1` and `data3` were sampled at $\Delta t = 0.001$ seconds and `data2` at $\Delta t = 0.002$ seconds.

The time base and thus the sampling interval in the data acquisition channel may not have been quite uniform. Together with the finite precision of the angle sensor, this has two important implications:

- Any ωs that you reconstruct using divided differences from the θ and time data may be inaccurate. You'll explore this in problem 1.

- Nonuniform sampling violates the conditions of the Takens theorem, so any attractors constructed via embeddings of these data are *not* true diffeomorphic copies of any attractor that may exist in the system...but they're pretty close. We can mitigate this by using embedding intervals that are much larger than the experimental sampling interval (or by interpolation, if we knew exactly how far off our samping interval was; see the optional reading listed above for details).

1. Write a program that steps through a data file, constructs an ω for each point using divided differences — first-order forward is good enough, but you may use something smarter if you want — and plots the data in state-space form, with θ mod 2π.

Note that if data are *oversampled* — that is, if the sampling rate is much faster than the device's dynamics — you have to be a little careful about what points you plug into the divided difference formulae. If the sampling rate is *slower* than the dynamics, on the other hand, you've missed some of the behavior.

Apply this program to `data1` and turn in a plot. Since the drive is off, this plot *should* be a clean spiral (why?). Please comment on what it really looks like, as well as on possible causes for this.

2. Write a program that steps through a data file and *embeds* the θ data, producing the corresponding trajectory in *reconstruction space*. This program should take a time interval τ, a dimension m, and indices j, k of a pair of axes on which to plot the results. It should produce a list of m-vectors (points in reconstruction space) each of whose i^{th} element is $\theta(t + i\tau)$ for $i = 0..m-1$. Finally, for each m-vector, it should plot the j^{th} element against the k^{th} element, both mod 2π.

> *Aside:* τ is usually an integer multiple of the sampling interval Δt in the data set; if it isn't, interpolation may be called for. (This is not an issue in this problem set.) In cases like this, people often sidestep the issue and just use the data point that is closest to the sampling interval. Sometimes, they do a linear interpolation between the points on opposite sides of the interval boundary. Note that $\theta(t)$ — the first coordinate of the reconstruction-space point — should *always* be a real data point. Again, see the optional reading listed above for details.

(a) Run your embedding program on the `data2` set with $\tau = 0.15sec$ and $m = 7$. Plot the zeroth element of the reconstructed state vector — $\theta(t)$ — on the vertical axis and the second ($\theta(t + 0.3)$, here) on the horizontal axis (i.e., $j = 0$ and $k = 2$). What kind of attractor is this? Turn in a copy of the plot.

(b) Now play with the `data3` set: use $m = 7$ and start with $\tau = 0.01sec$, then raise τ to 1.5, checking at least three intermediate points along the way — e.g., $\tau = 0.01,\ 0.15,\ 0.5,\ 1,\ 1.5sec$. What kind of attractor is this? Describe and explain the effects of the different τs and turn in one or two interesting plots — of $\theta(t)$ against $\theta(t + 5\tau)$ this time — that back up your explanations.

3. *[Thought experiment]* (a) In all of these problems, we used $m = 7$ whether or not the drive was on. What requirements does the Takens theorem place on m for a successful embedding of the *driven* pendulum? What about the *undriven* pendulum?

(b) What do you think would happen to the reconstructed trajectory — not just your picture, but the full trajectory — in part (b) of problem 2 if you had used $m = 2$ or $m = 25$? (one or two sentences only, please).

(c) What do you think would happen to the reconstructed trajectory in part (a) of problem 2 if you had used $\tau = 10^{-16}$ — which would require much more frequent sampling, obviously — or $\tau = 10^{6}$? (one or two sentences only, please). What would your pictures look like?