Visual Data Explorer

Chase Tisagh

Dynamical systems modeling often produces large sets of (multivariate) data, so modern computer algebra systems have a good visualization tools. Data visualization plays an important role in data analysis - visual inspection can help to reveal patterns that would be computationally rather difficult to reveal. The problem becomes especially involved when one deals with multivariate data. Unfortunately, not all computer algebra systems have good interactivity tools, and analyzing structures algorithmically is not easy. When we encountered this problem, it was decided to make a system for interactive exploration of multivariate data.

The system is designed to allow the user to find relations between two projections of a large multivariate dataset. Two selected variables are displayed on a scatter plot (left plot) with another two selected variables displayed on another scatter plot (right). The system can be scaled horizontally to produce as many plots as required from the same initial dataset. Points can then be selected on any plot and corresponding points will be selected automatically on the other plots. Figure 1 gives an example of such visualization and selection. Selected cases can be saved into a file for further analysis.

This system was initially developed using Microsoft Visual Studio IDE with Python and some of its supporting libraries designed for data analysis and visualization. Later, the system was migrated to Jupyter Notebooks in order to fully utilize the power of PyViz libraries specifically developed for this application. The main libraries imported were: Pandas; an open-source Python library designed for data science applications, and Holoviews with Bokeh on the backend; a Python library designed specifically for interactive data visualization using web browsers through JavaScript. Also Tkinter; to create a graphical user interface and accept user input, and NumPy; to effectively and efficiently perform mathematical operations if needed.

Author thanks for support Dr. C. V. Rao, Dean of Students, St. George's University, True Blue, Grenada, W. I. and The Windward Islands Research and Education Foundation, True Blue, Grenada, W. I.

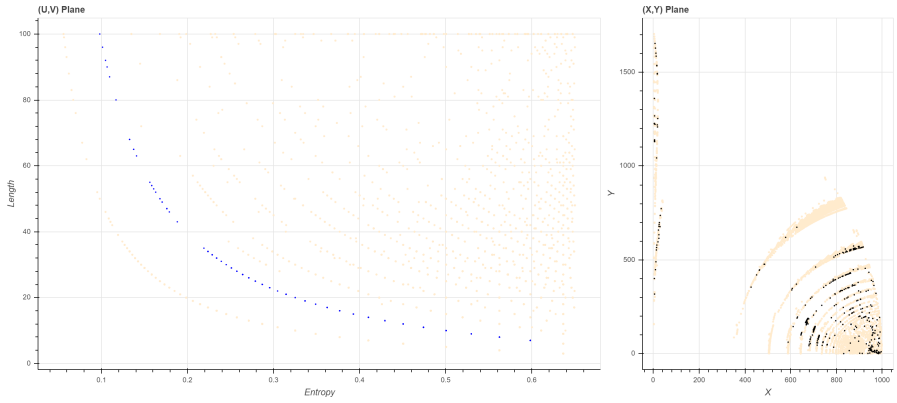

FIGURE 1. Left: One curve observed in the (U, V) projection is selected. Right: corresponding projection in another set of coordinates.

Chase Tisagh
Dept. of Computers & Technology, SAS
St.George's University
St.George's, Grenada, West Indies
e-mail: `CHATIS@sgu.edu`