Recent experiments involving the training of artificial neural networks with multiple layers, sometimes referred to as deep learning, have demonstrated the ability to automatically identify features that are critical to solving complex pattern classification tasks, such as speech recognition. Similar to speech, atmospheric data sets often consist of multiple time series with unknown, complex interrelationships. In this project we seek to explore what kind of interrelationships can be discovered in climate data by applying the framework of artificial neural networks. As a first application we look at establishing relationships between top of atmosphere radiative flux and air/surface temperatures. This is an important application, since a thorough understanding of those relationships is essential for understanding the effect of CO2-induced warming on the Earth's energy balance and future climate.

Objectives

1. Find low-dimensional (3) representation of each sample to study patterns among the 6 variables across spatial locations.

The key idea is to train neural networks from data and then study the network properties as function of location and field.

Principal Components Analysis

Using first 3 components of \( V \)

\[ X = \text{USV}^t = \text{svd}(X) \]

Approximation error (RMS) is 77.9

Neural Network as Autoencoder

Trained to minimize error in approximation (scaled conjugate gradient method)

Approximation error (RMS) is 23.0

Multiple Neural Networks as Autoencoders

Networks with single representation unit trained with approach akin to Gram-Schmidt.

Approach error (RMS) is 17.2

Discussion and Future Work

The example and initial results provided here illustrate the basic approach we plan to take, namely to train neural networks from data and to then study the network properties to discover interesting patterns in the data. The displays of weight magnitudes reveal subsets of locations and fields that are positively or negatively correlated in the set of data samples. We expect that further analysis of the network properties will reveal nonlinear combinations of key locations and fields that will facilitate an understanding of spatial and temporal variations in the data. We have only just started to implement these ideas, so the initial results presented here serve as an illustration of the general process.

NASA CERES and MERRA Data

- \( lw \) – long wave flux at TOA
- \( st \) – skin temperature
- \( si \) – solar insolation at TOA
- \( t1 \) – air temperature 500 hPa
- \( t2 \) – air temperature 50 hPa

Each sample composed of 7 x 18 x 6 = 756 variables.

Daily data from March 1, 2000 – December 31, 2013 (5,054 days)

Two primary data sources:

- \( USV \)
- \( X \)

Approximation error (RMS) is 23.0

\[ V_1, V_2, V_3 \]

Components of eigenvectors as function of location and field. Small magnitude values are not shown.

Positive

Negative

Input

First Layer

Representation Units

First Layer Units

Unit

Unit

Unit

Approximation of input

Approximation of input

First Layer Units

Representation Units

First Layer Units

Representation Units

Approximation of input

Better:

Only first unit represents the annual cycle.

Weights of individual units in first layer as function of location and field. Small magnitude values are not shown.

Below figures show weights of individual units in first layer of second and third networks.