Finding Higher Order Interactions Using Local Corex

Thomas Kerby, Kevin Moon, & Greg Ver Steeg

Department of Mathematics and Statistics
Utah State University

April 7, 2022

What are HOI's and Why do we care?

What: Higher Order Interactions (HOI's) are interactions that occur in the presence of three or more variables.

Why:

- Simplifies problem scope
- Increases model interpretability
- Improves model predictions

Credit: https://www.flickr.com/photos/cupcakexcult/3766846693/

How do we find HOI's?

This is an unsolved problem.

- Pair-wise interactions grow on the order of n^2
- HOI's lie in a search space far too large to search by brute force when n is large

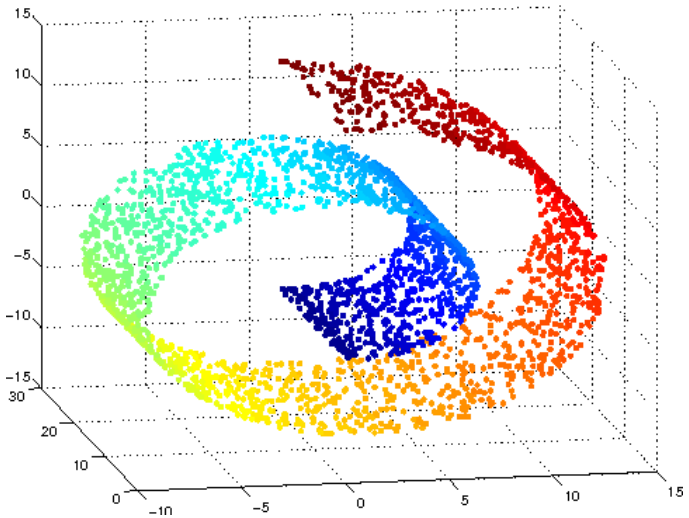
We need smart ways to detect HOI's without searching the entire search space

Changing HOI's

HOI's often change depending on their "location" in the data manifold.

Why?

- Treatment Effects
- Multiple Observation Types

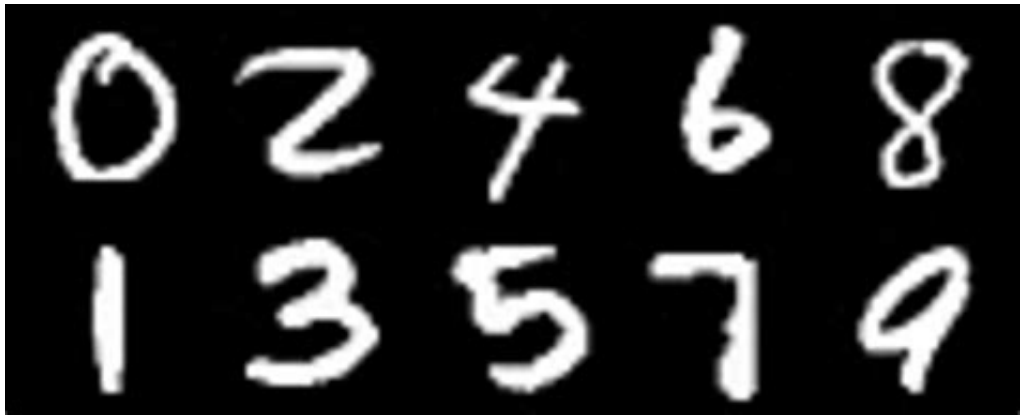


Changing HOI's - Continued

All other methods assume that HOI's are consistent

Example:

- Pixel relationships depend on the number
- Multiple groups may exist for a single class



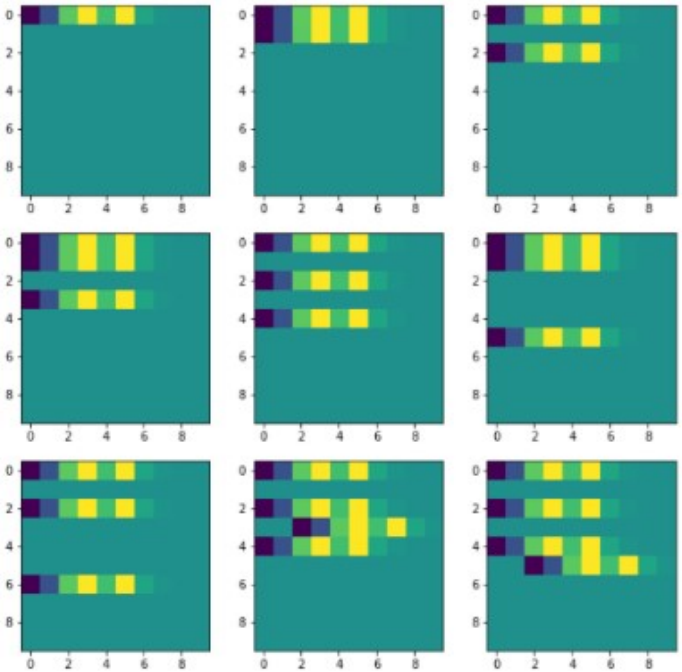
Linear Corex

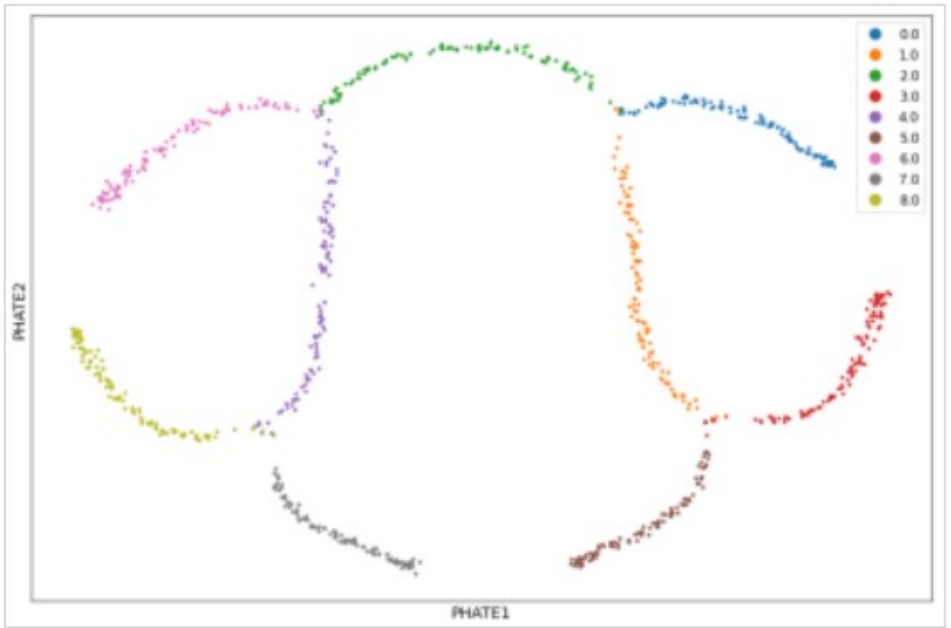
- Creates a Latent Factor Representation of the data
- $min_W TC(X|z) + TC(z)$, where $z = Wx$
- Latent factors are added such that they maximally explain your dependence or Total Correlation
- Weaknesses
  - Linear Method
  - Doesn't give inference into the nature of variable relationships
  - Struggles to capture full groupings of related variables when variable interactions change along the data manifold

Local Corex Overview

- Partition data on the data manifold
- Run Linear Corex on each partition to transform it
- Take average observation in partition and perturb the value of chosen variable. Repeat many times
- Untransform perturbed observations
- Look at which variables have the most variance

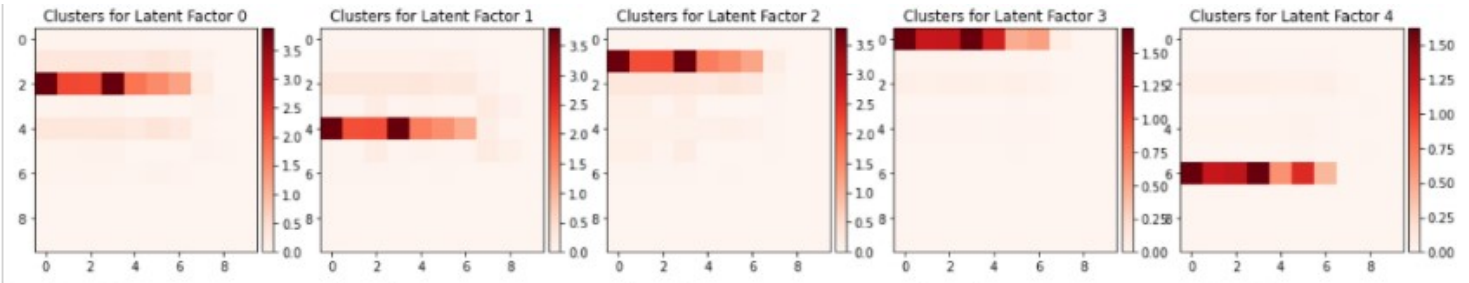
Synthetic Tree Data



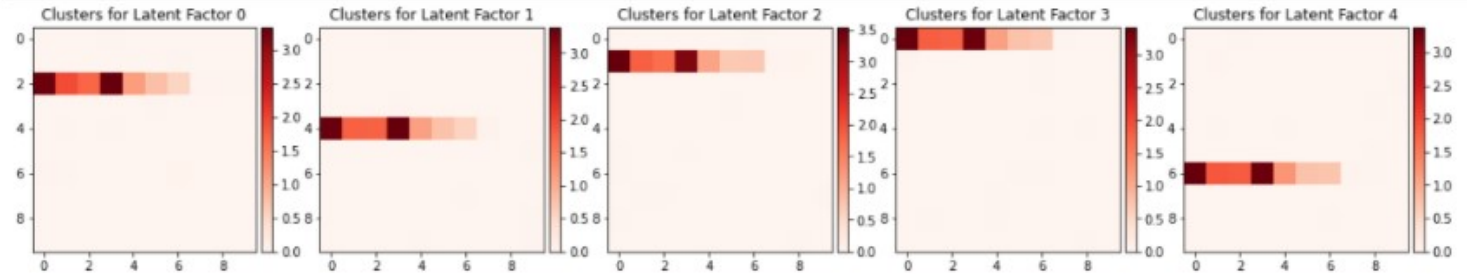


Linear vs Local Corex

Linear Corex

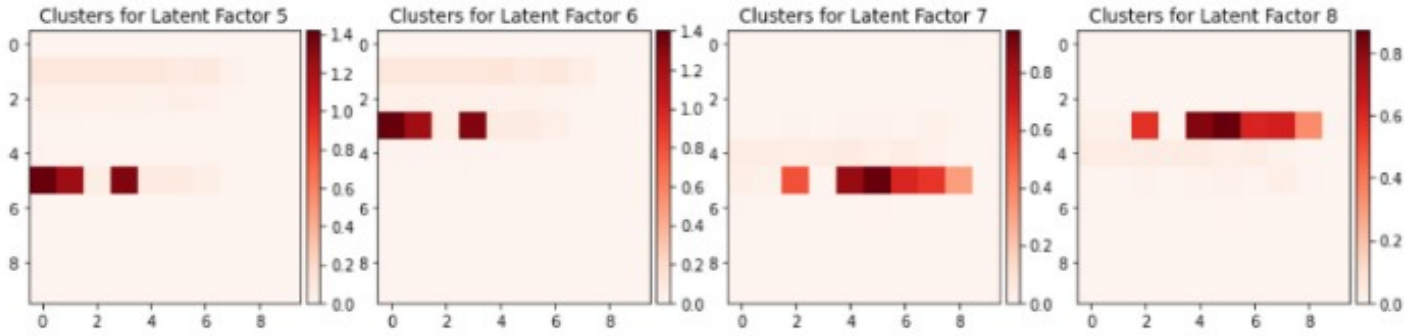


Local Corex

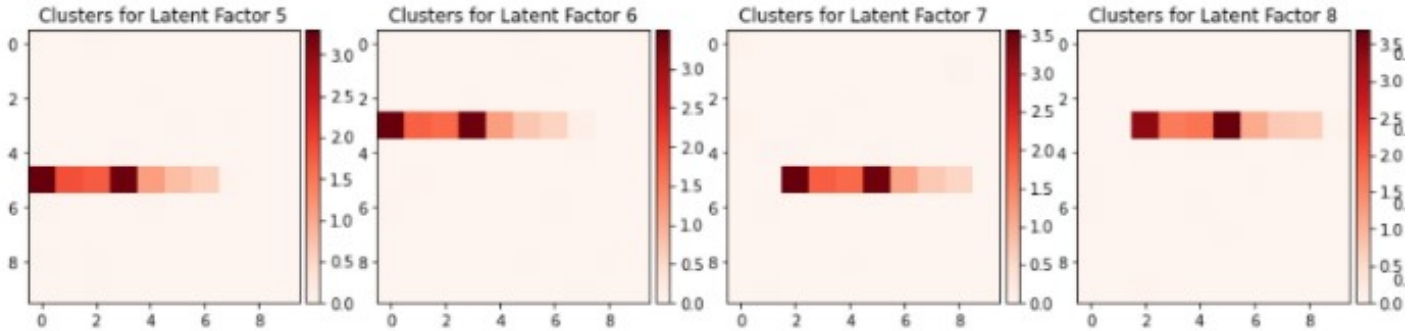


Linear vs Local Corex - Continued

Linear Corex



Local Corex



Summary

- HOI's are hard to find - especially when the data changes
- Linear Corex efficiently captures HOI's when the data doesn't change
- Local Corex is able to efficiently capture HOI's even when the data changes

Future Work

- Apply to more data
  - Synthetic
  - Real World
- Determine what kinds of HOI's can and cannot be found using Local Corex
- Compare against more algorithms
- Create a nonlinear version