Spectral Clustering

Renato R. Maaliw III, *DIT*
College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines

Spectral Clustering

- a way to group things that are similar to each other
- it uses a concept in linear algebra called eigenvectors and eigenvalues to simplify information and make it easier to work with

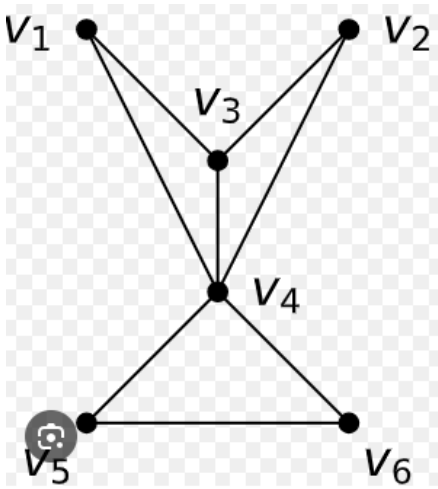
Once simplified, spectral clustering can group similar things together.

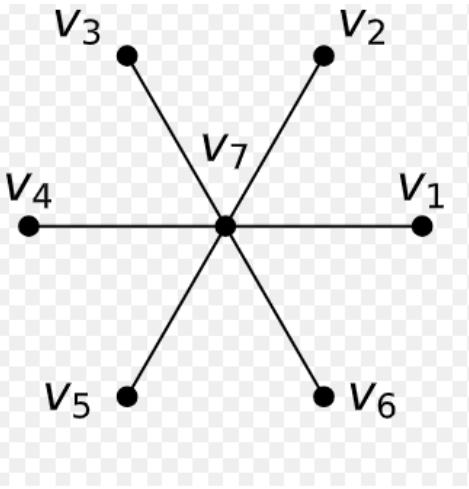
Cognate/Professional Electives

For example, if you were using it to understand a social network, it could help you identify different groups of people who are connected to each other

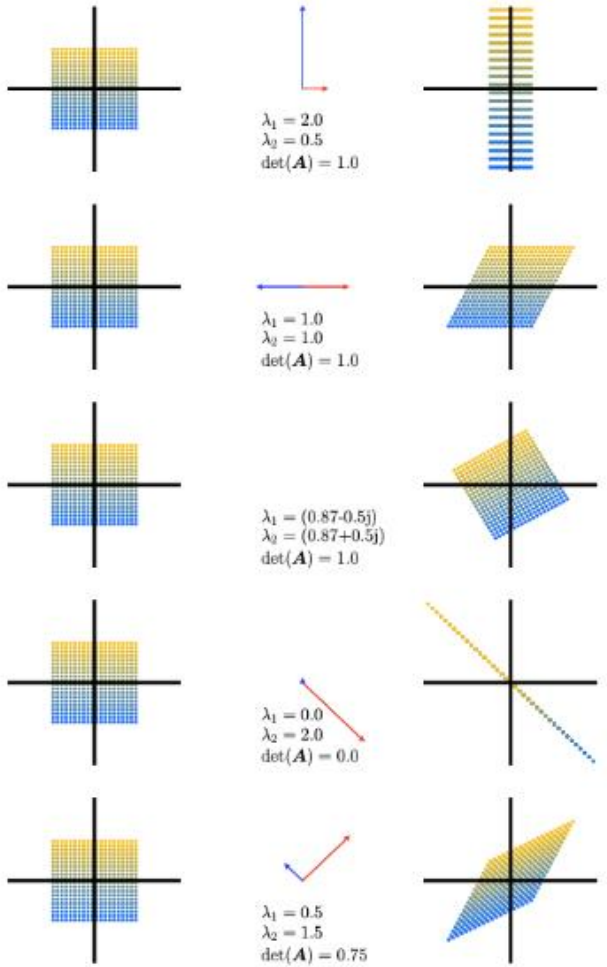
It relies on matrix (graph laplacian) allowing to use it for computing clusters.

We compute the graph laplacian using basic properties of graphs.





Simply, it turns your data in a network (graph) of similarities, then uses a mathematical “listening” trick (eigenvectors) to discover natural splits



Why do wee need Spectral Clustering?

- real data rarely forms neat circles: such as interlocking crescents, spirals, thin bands, or manifold-shaped clouds

Why do wee need Spectral Clustering?

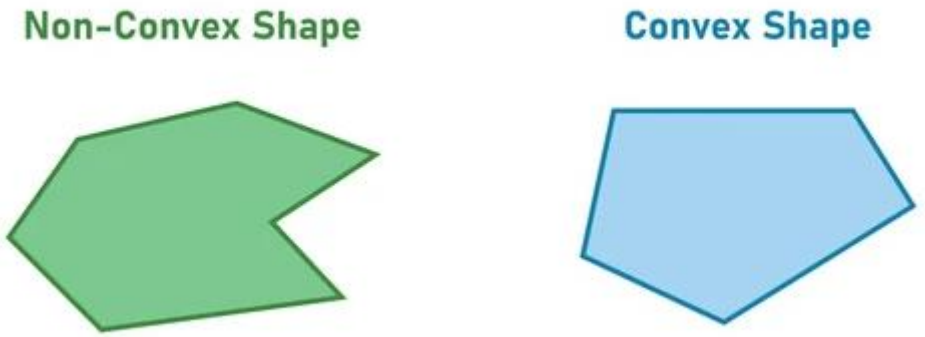
- distance-only methods (like K-means) can cut through a crescent, mixing different regions

Why do wee need Spectral Clustering?

- spectral clustering respects local neighborhoods and the geometry of the data

Where Popular Methods Struggle

- K-Means: assumes round, equally sized clusters; struggles with non-convex shapes



Where Popular Methods Struggle

- Hierarchical (linkage): sensitive to linkage choice; may chain across gaps

Where Popular Methods Struggle

- Gaussian Mixture Model (GMM): assumes cluster look like (possibly rotated) ellipses; complex shapes need many components. Has to pass the normality assumptions.

Spectral clustering: uses connectivity rather than global shape assumptions

Thank you very much for listening.