1. Pairwise Display and the PairViz R package

Wayne Oldford based on joint work with Catherine Hurley

Tutorial B1
The problem

- Can we automatically, yet meaningfully, layout complex statistical displays?

- Can we navigate high dimensional structure in a useful yet controlled way?

- Answer: Yes
graph theory provides framework, Statistics adds meaning.
Radial axes

7 variates: 
   A, B, C, D, E, F, G
Radial axes

7 variates:
    A, B, C, D, E, F, G

Arrange axes as equiangular radii:
Radial axes

Length of each radius is proportional to (scaled) variate value for that case.

Have a “star-shaped” glyph for each case.
Radial axes:

- Compare cases by shape of glyph,
- Compare cases by shape of glyph, here 9 cases in 7 dimensions
- Visually cluster high dimensional data by shape:
  \{7,8,9,1\} \{2,3\} \{4\} \{5,6\} ?
Radial axes:

- Compare cases by shape of glyph, here 9 cases in 7 dimensions
- Visually cluster high dimensional data by shape:
  \{7,8,9,1\} {2,3} {4} {5,6} \\
- What if the variables were assigned in a different order?
Radial axes: order effect

Dataset order H0

\{7,8,9,1\} \{2,3\} \{4\} \{5,6\} ?

Order H1

\{7,8,9,1\} \{2,3\} \{4\} \{5\} \{6\} ?

Friday, November 25, 2011
Radial axes: order effect

Dataset order H0

\{7,8,9,1\} \{2,3\} \{4\} \{5,6\}  

Order H1

\{7,8,9,1\} \{2,3\} \{4\} \{5\} \{6\}  

Order H2

\{1,4\} \{2,3\} \{5,6\} \{7,8,9\}  

Friday, November 25, 2011
Radial axes: order effect
... different orders = different visual effect

Dataset order H0

Order H1

Order H2

Order H3
Radial axes: reduced order effect

A variate ordering = Hamiltonian path

Friday, November 25, 2011
Radial axes: reduced order effect

A variate ordering = Hamiltonian cycle or tour
Radial axes: reduced order effect

A variate ordering = Hamiltonian cycle or tour

There are lots of Hamiltonians to choose ... A Hamiltonian visits every node on the complete graph $K_7$
Radial axes: reduced order effect

A variate ordering = Hamiltonian cycle or tour

There are lots of Hamiltonians to choose ...

A Hamiltonian visits every node on the complete graph $K_7$

An Eulerian Tour visits every edge once (nodes possibly many times)
Radial axes: reduced order effect

A variate ordering = Hamiltonian cycle or tour

There are lots of Hamiltonians to choose ...

An Eulerian Tour visits every edge once (nodes possibly many times)

Non-(edge)-intersecting Hamiltonian tours whose union is the entire graph, is called a Hamiltonian Decomposition.
Radial axes: reduced order effect

A variate ordering = Hamiltonian cycle or tour

There are lots of Hamiltonians to choose ...

A Hamiltonian visits every node on the complete graph $K_7$

An Eulerian Tour visits every edge once (nodes possibly many times)

Non-(edge)-intersecting Hamiltonian tours whose union is the entire graph, is called a Hamiltonian Decomposition.
Radial axes: reduced order effect

A variate ordering = Hamiltonian cycle or tour

There are lots of Hamiltonians to choose ...

A Hamiltonian visits every node on the complete graph $K_7$

An Eulerian Tour visits every edge once (nodes possibly many times)

Non-(edge)-intersecting Hamiltonian tours whose union is the entire graph, is called a Hamiltonian Decomposition.
A variate ordering = Hamiltonian cycle or tour

There are lots of Hamiltonians to choose ...

A Hamiltonian visits every node on the complete graph $K_7$

An Eulerian Tour visits every edge once (nodes possibly many times)

Non-(edge)-intersecting Hamiltonian tours whose union is the entire graph, is called a Hamiltonian Decomposition.
Radial axes: reduced order effect

An Eulerian

A Hamiltonian
Radial axes: reduced order effect

A Hamiltonian decomposition
Radial axes: reduced order effect

A Hamiltonian decomposition

- Which when assembled form an Eulerian cycle composed of Hamiltonians
Radial axes: reduced order effect

- Which when assembled form an Eulerian cycle composed of Hamiltonians
- Could build a glyph from these cycles (21 radii instead of 7)
Radial axes: reduced order effect

A Hamiltonian decomposition

Hamiltonian decomp, H1:H2:H3

- Could build a glyph from these cycles (21 radii instead of 7)
Radial axes glyphs

A Hamiltonian decomposition

Hamiltonian decomp, H1:H2:H3
{1,4} {2,3} {5,6} {7,8,9}
Radial axes glyphs

A Hamiltonian decomposition

Hamiltonian decomp, H1:H2:H3

\{1,4\} \{2,3\} \{5,6\} \{7,8,9\}

Clustering dendrogram
Radial axes glyphs

A Greedy Eulerian (maximizing pairwise correlation)

Eulerian order

Clustered axes glyphs

Clustering dendrogram

Height

Friday, November 25, 2011
Radial axes glyphs

A Greedy Eulerian (maximizing pairwise correlation)

First two principal components

Clustering dendrogram

All pairs (Greedy Eulerian or Hamiltonian decomposition) reduce the effect of variable pair patterns, making star glyphs more reliable.
Radial axes glyphs

A Greedy Eulerian (maximizing pairwise correlation)

First two principal components

Clustering dendrogram

All pairs (Greedy Eulerian or Hamiltonian decomposition) reduce the effect of variable pair patterns, making star glyphs more reliable.
Radial axes

• Hamiltonian decomposition
  • all pairs of variates appear so no one pair dominates
  • divides glyph into sectors with all variates appearing once in each sector
  • ameliorates well known order effect
  • there are many Hamiltonian decompositions for a complete graph
    • For $K_7$, two generator decomps.
  • could choose Hamiltonian to maximize some measure on sum of edges (TSP)
Radial axes

• Eulerian
  • all pairs of variates appear
  • can order pairs (greedy Eulerian)
  • could choose fewer radial axes (early pairs are emphasized).
  • lots of Eulerians
    (e.g. for $K_7$: $129,976,320$;
    $K_{21}$: $> 3.5 \times 10^{184}$)
More abstractly ...

- Visiting all nodes once would give all objects in some order, a Hamiltonian.

- Visiting all edges once would give all object pairs in some order, an Eulerian.

- Visiting all edges via a Hamiltonian Decomposition gives all pairs of objects and distributes the objects so that all appear in a block.

- There is still considerable choice on the order of the nodes in a walk within these constraints; possibly accommodate with weights on edges.
More abstractly ...

Order of objects = a walk or tour on the complete graph

- Visiting all nodes once would give all objects in some order, a Hamiltonian.

- Visiting all edges once would give all object pairs in some order, an Eulerian.

- Visiting all edges via a Hamiltonian Decomposition gives all pairs of objects and distributes the objects so that all appear in a block.

- There is still considerable choice on the order of the nodes in a walk within these constraints; possibly accommodate with weights on edges.
PairViz demo
Summary

- Graph theory has much to offer
  - map “objects” to nodes
  - map “transitions” or “comparison” to edges
  - add statistically meaningful weights to edges
  - use weights to guide the visual search
  - use Hamiltonians, Eulerians to reduce unintended effects.

- Lots more to explore ... constructions, decompositions, weights, applications, ...
Thank you
Thank you

Questions?

御清聴ありがとうございますました

質問はありますか？
Hurley & Oldford:

Pairwise display of high dimensional information via Eulerian tours and Hamiltonian decompositions (*JCGS*, 2010)

Eulerian tour algorithms for data visualization and the **PairViz** package (*Comp Stats* 2011)

**PairViz** R package ... available on CRAN.