California —— 58 counties —— 24626 precincts

Red : Trump

Blue : Clinton

white : equal

darker color ⟶ stronger voting preferences

Fig. 1

Goal: Find "island" which a region votes differently from the areas that surround it.

§ Construct Simplicial Complexes

Only for blue / red area

① V-R complex.

X: data-set, $\varepsilon > 0$

$VR_\varepsilon(X) = \{\sigma \subseteq X : \forall x, y \in \sigma, d(x,y) \le \varepsilon\}$

drawback: worst case has $2^n - 1$ simplices and $n-1$ dim for n points.

Thus we compute alpha complex when precincts > 150.

② alpha complex

X: data set $\varepsilon > 0$ $(V_x)_{x \in X}$: Voronoi diagram.

$A_\varepsilon(X) = \{\sigma \subseteq X : \forall x_i \in \sigma, \bigcap_i (V_{x_i} \cap B(x_i, \varepsilon)) \neq \emptyset\}$, V_{x_i} is cell.

Advantage: dim of simplices at most 2, since embedded into diagram

Remark: Both ① and ② are distance based method. simplicial complexes cannot reflect the underlying geographical map.

Fig 2, Fig 3.

③ Adjacency complex.

Def: two precincts are "queen adjacent" if they touch

at any two points, including corners.

Define $\delta_{b,r}(p) = |V_b(p) - V_r(p)| / |V_b(p) + V_r(p)|$, where $V_b(p)$ is the number of blue votes in a precinct p. and $V_r(p)$ is the number of red votes.

Now we may consider the precincts that Trump (Red) won. First consider $\delta_{b,r}(p) > 0.95$, that is highly prefer Trump. Then decrease $\delta_{b,r}(p) > 0.90 \cdots$ At each step, we construct an edge if two precincts adjacent; we construct a 2-simplex between any three vertices that are all pairwise adjacent.

Thus we get 2D filtered simplicial complex.

Fig. 4

Advantage: get information about precincts

Disadvantage: still associate a single point to each precinct polygon

④ Level-Set Complex

Let M be 2-mfd that consists of the collection of all of a country's precincts that voted for the same candidate. $\Gamma = \partial M$.

Consider $\phi(\vec{x}, t) : \mathbb{R}^2 \times \mathbb{R} \longrightarrow \mathbb{R}$.

$\phi(\vec{x}, 0)$ is the signed distance function from $\vec{x}$ to the boundary Γ.

0-level set of $\phi(\vec{x}, 0)$ is precisely Γ.

Evolve ϕ by $\frac{\partial}{\partial t}\phi = |\nabla \phi|$.

In each time T, consider the 0-level set Γ_T.

Figure 5. 6.

Take triangulation, get filtered simplicial complex.

(5th pixel → 1 vertex) Figure 7

Advantage: (i) gives an explicit triangulation of a geographical map that does not depend on how we assign precincts to points.

(ii) Every hole is a feature that is born at time 0. Noise occur later due to complicated shape of precincts.

Disadvantage: missing small island when triangulation.

⑤ Comparing.

Problem { Scaling: island occur at a variety of distance scales
Contiguity: physical distance does not correspond to geographical connectedness.

Table 1

§ Computation Results.

Sizes: Table 2.

① Adjacency and level-set complexes do not scale in size as rapidly as the VR complex (e.g. Merced, Napa)

② Number of precincts is rather small, a level set complex can still be rather large (e.g. Lake, Lassen).

Computation Time: Table 3

① Adjacency and level-set complex are significantly faster than VR complex, slight slower than alpha complex.

Barcodes & Feature Maps

Feature map. Fig. 8 circle ⟶ H_1 generator
darker ⟶ persistence longer

Example 1: Red Precincts in Tulare County (Figure 9)

Background: 10 precincts rural urban Republican (Red)

Goal: detect blue island

Figure 10.

Example 2: Blue Precincts in Imperial County. (Fig. 11)
NOT evident where there may be holes.