spatial auto-correlation (pysal)

adam.okulicz.kozaryn@gmail.com

this version: Friday 22nd November, 2024   09:52

outline

intuition

spatial weights

using spatial weights: spatial auto-correlation

TODO: K-means, medians etc

outline

intuition

correlation and spatial auto-correlation

- everyone heard of correlation, right? what is it? examples?

- many things correlate positively; people in space, too:

  ○ obese with obese, smokers with smokers, etc
  ○ ppl you hang out with are like you–you're avg of your 5 or
    so closest ppl https://www.ncbi.nlm.nih.gov/pmc/articles/
    PMC3343772/pdf/nihms-216230.pdf p11 vis

  ○ https:
    //www.google.com/search?q=christakis+fowler+obesity

- http://www.thebigsort.com/maps.php

- hi-crime neighb next to hi-crime neighb

- poor states next to poor states (Miss, Alab, etc)

- poor countries cluster together, too: Africa, Lat Am, etc

the first law of geography (Waldo Tobler)

- "Everything is related to everything else,
  but near things are more related than distant things"
- almost always true!
- do you see this in your research?

outline

intuition

spatial weights

using spatial weights: spatial auto-correlation

TODO: K-means, medians etc

spatial weights–the first step

- the first step before producing spatial auto-corr
- is to produce spatial weights
  or spatially lag a variable

we will spatially lag a variable

- it's like time lagging a variable

  ```
  https://www.google.com/search?q=time++lag+
  variable&client=firefox-b-1-d&sca_esv=
  967cb4d33a83a1c2&udm=2&biw=1439&bih=1449&
  sxsrf=ADLYWIJ9CzRKGiuiQgClhLdRPammoCUzEg%
  3A17322855491166&ei=bZRAZ_PqCabdptQPrJSjyAM&
  ved=0ahUKEwiz1brGkvCJAxWmrokEHSzKCDkQ4dUDCA8&
  uact=5&oq=time++lag+variable&gs_lp=
  EgNpbWciEnRpbWUgIGxhZyB2YXJpYWJsZUgAUABYAHAAeAC
  sclient=img
  ```

  ○ time lag alows analysis of temporal precedence

  ○ eg corr/effect of unemployment last year on this year's

spatial lag

- spatially lagged var: want to know the relationship of a place to its neighbors
- spatially lagged variable is just
  ○ an average of values for its neighbors
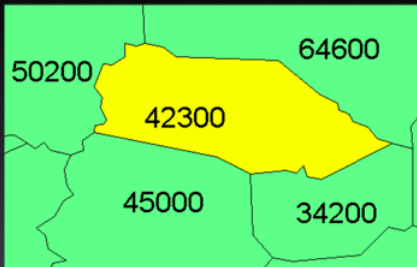
2 ways

- contiguity based (we'll just do these):
- neighbor of place A touches on place A
- distance based: neighbor of place A is within some distance of place A

2 types of contiguity weights

- usually just pick queen contiguity–neighbor is any place that neighbors our place
- at least must share a vertex, say North, North-East, etc
- can do rook: must share a border, not just vertex
- so *not* North-East

rook v queen

-
- Rook: only 2,4,6, 8; Queen: all (i.e. 1-8)

order of contiguity

- can be higher orders: neighbors of my neighbors are my neighbors
- we'll just do 1st order

outline

intuition

spatial weights

using spatial weights: spatial auto-correlation

TODO: K-means, medians etc

## positive v negative spatial autocorrelation

- note: <u>auto-</u>correlation
- correlate values of a var with values of the same var
- corr the variable with its spatial lag (avg for neighbors)
- positive if similar values next to each other
- negative if dissimilar values next to each other
- can already see in plain thematic maps

pos and neg

POSITIVE : Pattern of Similarity
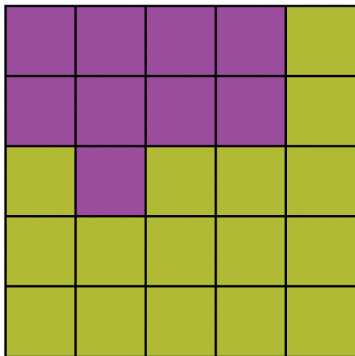
NEGATIVE : Pattern of Dissimilarity
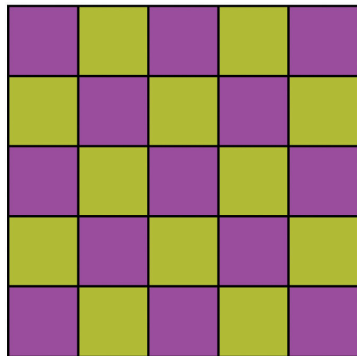
negative correlation is even more interesting

- uncommon: more interesting
- if a dog bites a man its no news,
  if a man bites dog, that's news
- eg see rich area in the middle of poverty

application: my paper about happiness in Europe

- https://theaok.github.io/docs/gesis3.pdf
○ see histogram and maps
- positive spatial autocorrelation
- clusters of happy and unhappy provinces
○ and they span across country boundaries
○ it is interesting to identify them and formally test

Moran's I

- most basic spatial stat
- just like regular corr: -1 to 1
- see Moran's I scatter in

  https://theaok.github.io/docs/gesis3.pdf

just a thematic map

- already see spatial auto-correlation in thematic maps
- just have a close look, and think about it, discuss in ps

so what?

- Moran's I and LISA help make sense of thematic maps
- they identify patterns, clusters, outliers
- very useful! eg is poverty concentrated?
- can already see from thematic maps but Moran's I and LISA add precision beyond eye-balling choloropleths
- always: don't forget about interpretation
- (don't ever show anything that you don't interpret)

so what?

- and it does matter where in the cluster one is located
- eg being poor in the middle of poverty may be better
- than being poor next to rich
- https://academic.oup.com/qje/article/120/3/963/1841496
- https://www.jstor.org/stable/10.1086/603534
- https:
  //journals.sagepub.com/doi/abs/10.1177/1078087418755515

outline

intuition

spatial weights

using spatial weights: spatial auto-correlation

TODO: K-means, medians etc

DEFINITELY DO NEXT TIME