Center for Spatially Integrated Social Science

CSISS Tools Project
Status and Plans Year 5

Luc Anselin
Spatial Analysis Laboratory
Dept. Agricultural and Consumer Economics
University of Illinois, Urbana-Champaign

Outline

- Highlights of Year 4
- Tools Web Site
- GeoDa
- Other Software Development
- GeoDa Demonstration

Highlights of Year 4

Team

- Director: Luc Anselin
  - software design, methods, algorithms
  - user's guide and tutorials
- Research Associates: Oleg Smirnov, Yongwook Kim
  - software engineering, algorithms
- Postdoctoral Visitor: Julie Le Gallo (Bordeaux)
  - methods, panel data spatial econometrics
- Graduate Research Assistants: Ibnu Syabri (Geography), Youngihn Kho (Computer Science)
  - software implementation

Software Development

- GeoDa
  - introduction to spatial data analysis (C++)
  - first release of beta (2/03)
  - approx 950 downloads so far, active use
- PySpace
  - spatial regression/spatial econometrics (Python)
  - prototype near complete
- OpenSpace
  - web-based spatial data analysis (Java)
  - prototype implemented

Supporting Materials

- Promote Understanding of Spatial Data Analysis Methods
  - Openspace Mailing List
    - +/- 250 subscribers
  - Tutorials
    - GeoDa User's Guide and 2 Tutorials
    - R, CrimeStat, Variowin

Supporting Materials (2)

- Facilitate Application of Methods
  - Sample Data Sets
    - ESRI shape file format
    - from 32 to 3085 observations
    - facilitate replication of published studies
  - Spatial Weights Archive
    - web based interactive interface
    - US states, US counties
    - counties in each state
    - GAL format (GeoDa, R spdep)

Supporting Materials (3)

- Illustration
  - web-based spatial data analysis
    - Java
  - using sample data sets
  - functionality
    - rate mapping and smoothing
    - outlier maps
    - Moran scatterplot
  - Anselin, Kim, Syabri (2004) Journal of Geographical Systems

Collaboratives

- GeoVista Studio
  - development of spatial data analytical modules (NCI, CDC)
- CrimeStat
  - assistance with regression/spatial regression implementation (NIJ)
- Rgeo
  - international collaborative to promote and consolidate spatial data analysis in R

Other

- Advances in Spatial Econometrics
  - Anselin, Florax, Rey (eds.)
  - CSISS “best practice” (to publisher 12/03)
- Tools Workshop Special Journal Issues (04/05)
  - Computers and GeoSciences
  - Geographical Analysis
- Spatial Externalities Special Journal Issue
  - International Regional Science Review 26 (2), 2003

Tools Web Site

http://sal.agecon.uiuc.edu/csiss/

GeoDa

Goal

- An introduction to spatial data analysis
  - visual, interactive, user-friendly
  - path from basic visualization EDA and ESDA to spatial regression
  - for non (or not yet) GIS users
- Free
  - “official” release
    - http://sal.agecon.uiuc.edu/csiss/Geoda.html
  - latest (bleeding edge)
    - http://sal.agecon.uiuc.edu/stuff/fixes/
- Current Version 0.95

Mapping Functionality

- Choropleth maps
  - quantile, percentile, standard deviational
  - outlier maps: box map
  - rate smoothing
    - EB smoothed rates, spatial smoother
- Map Movie
  - animation
- Cartogram
- Conditional Maps (multivariate)

Statistical Graphics/EDA

- Linking and Brushing
  - dynamic linking and brushing of all maps and charts
- Univariate
  - histogram, box plot
- Bivariate
  - scatter plot (correlation)
- Multivariate
  - parallel coordinate plot
  - 3-D visualization

Spatial Autocorrelation

- Moran Scatterplot
  - Moran's I with permutation test
  - bivariate Moran's I
  - variance instability adjustment for rates (EB)
- Local Moran
  - LISA map, significance map, box plot of Local Morans
  - univariate, bivariate, EB adjusted
  - save results to table

Spatial Regression

- OLS with diagnostics
  - tests for spatial autocorrelation
- Maximum Likelihood
  - lag and error model
  - implements large data set algorithms
    - e.g. spatial hedonic regression with N = 340,000
- Save Results
  - predicted values and residuals for mapping and data export

Utilities

- Spatial Weights
  - contiguity and distance based
  - for points and polygons
  - spatial weights characteristics
- Queries and Computations
  - table
- Spatial Data Manipulations
  - centroids, Thiessen polygons, regular grids
  - data input/output
    - ascii to shape file and shape file to ascii

Plans for Year 5

- Release of V1.0 by Summer 2004
  - complete spatial regression functionality
  - finalize visualization
  - exploratory variography
  - (spatial data analysis on networks)
- User's Guide
  - help files
- Community Building
  - mailing list, workshops
  - supporting materials

Other Software Development Projects

PySpace

- Prototypes
  - linear regression model
    - ML/IV/GMM estimation
    - full slate of diagnostics
    - no sparse weights yet
  - panel spatial econometrics
    - pooled cross-section time series
      - OLS with diagnostics, lag and error models
    - spatial SUR

Analysis of Large Data Sets

- Asymptotic Variance Matrix
  - ML lag and error estimation
  - requires inverse of N by N matrix
    - new algorithm (Smirnov)
    - tested on N > 1 million
- Parallel Computation
  - parallelizing ML estimation
    - new algorthim (Smirnov and Anselin)
    - prototype implemented in Java

Plans for Year 5

- Release of PySpace
  - target date summer 2004
  - consolidation of current rewrite
  - move to SourceForge
- PySpace User's Guide
- ESRI ArcGIS 9.0 Extension
  - Experimental

GeoDa Demonstration