Spatial Statistics and Spatial Econometrics (ECO 324/524)
Mid-Semester Exam

Maximum marks = 50

Instructions: This is an open-notes; open-book exam. Use of internet during the exam is not allowed. Total time allotted is 1.5 hours: 45 minutes for the theory section; 45 minutes for the empirical section. Theory section must be attempted before the empirical section.

Theory Portion (25 points)

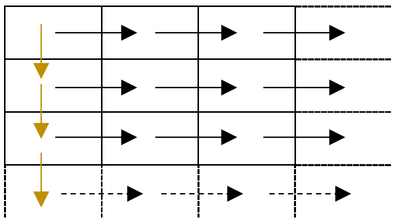
Q1.

- Consider that a large landmass is divided into an *n* x *n* lattice (i.e., *n* rows and *n* columns) where each one-acre cell represents a land parcel. Denote the land value of each parcel, indexed by its location and measured in rupees, as $V_{\mathbf{ID}}$. Clearly, $\mathbf{ID} \in R^2$ and $\mathbf{ID} = \{x, y\}$ where *x* can be taken to represent row# and *y* to represent column#.
- Now assume, there exist localized spatial network economies in land values. Specifically,

i. In every row, the market value of each parcel in rows is (partially) dependent on the value of the parcel on its immediate left.

ii. In the first column, the market value of each parcel in rows is (partially) dependent on the value of the parcel that is located above it.

Arrows in the following schematic diagram portrays spatial dependence along the lattice



- Define a spatial weights matrix, *W*: a square-matrix and a symmetric matrix where each entity $w_{i,j}$ = 1 if the land value of parcel *i* is spatially dependent on the value of parcel *j*, and 0 otherwise. Of course, the diagonal elements of *W* are all 0s.

  (a) What is the size of the spatial weights matrix, *W*. Determine the total number of unique elements in *W*.

  (b) Assume *n* = 10. Assign a left-to-right numbering scheme to land parcels in order to map vector **ID** to a scalar *i* such that *i* = 1 for $\mathbf{ID} = \{1,1\}$; *i* = 10 for $\mathbf{ID} = \{1,10\}$; *i* = 11 for $\mathbf{ID} = \{2,1\}$; *i* = 21 for $\mathbf{ID} = \{3,1\}$; and so on.

  Determine all the elements of row#5 of the weights matrix. What are the elements of column#5 of the weights matrix? How do they differ from row elements and why? Would your answers to the above queries be different from those for row#1 and column#1? If no, why? If yes, how?

- Denote vector $\vec{V} = [V_1, V_2, V_3, ...]'$ such that it contains land values of parcels in the lattice. The *spatial-lag* vector $\vec{L} = W\vec{V}$ represents spatial dependence among land parcels such that each row *i* in $\vec{L}$ contains values of those parcel(s) that impact V_i due to the aforementioned network effects in the land economy.

(c) Assuming n = 10 and the parcel numbering scheme in part (b), determine the following elements of the matrix $\vec{L}$: (1, 1); (51, 1); and (100, 1).

(d) Assume that $V_1 = V_{\{1,1\}} = V_0 \sim N\left(\mu, \sigma_V^2\right)$ and $\varepsilon_{\mathbf{ID}}$ is *iid* random variable with mean 0 and variance σ_ε^2. Determine which of the following random functions representing land value spatial dependence is second-order stationary?

i. $\vec{V} = k + \rho W\vec{V} + \vec{\varepsilon}$, where k is a real-valued constant and $|\rho| < 1$.

ii. $\vec{V} = ki + \vec{\varepsilon}$, where k is a real-valued constant.

iii. $\vec{V} = ki + \rho W\vec{V} + \vec{\varepsilon}$, where k is a real-valued constant and $|\rho| < 1$.

In the above cases where the processes are second-order stationary, is it also second order stationary.

(e) Now assume that besides the spatial spillovers, land values are also driven by each parcel's soil quality, say Q_i. So a more comprehensive model for the land parcels is given as

$$\vec{V} = \rho W\vec{V} + Q\beta + \vec{\varepsilon} \qquad (1)$$

- Write down is the dimension of each element of equation (1).
- Show that we can re-write equation (1) in the following form

$$\vec{V} = \left(I - \rho W\right)^{-1} Q\beta + \left(I - \rho W\right)^{-1} \vec{\varepsilon} \qquad (2)$$

where I represents an identity matrix.

- (<u>*Bonus question*</u> = 5 points) Say you know that if the lattice is large-enough, $\left(I - \rho W\right)^{-1} \approx I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \ldots$

Can you envision disentangling the parcel-level and the neighbourhood-level effects of soil quality on land values using model (1). Explain.

Empirical Portion (25 points)

Q2.

You are provided a 23x16 lattice containing an observed sample of ore-density values, measured in tonnes per cubic meter.

(a) Provide the descriptive statistics of the provided data. Comment on the empirical PDF and CDF of these data.
(b) Examine spatial dependence and spatial heterogeneity in these data.
(c) Identify outliers in the ore-density data. (Hint: Besides your general understanding of outlier, refer to question 2 of Homework 3).