Spatial Statistics and Spatial Econometrics

Prof. Gaurav Arora
Department of Social Sciences and Humanities
Indraprastha Institute of Information Technology, Delhi

Lecture - 18A
Spatial Weights Matrix

Hello everyone. Welcome back to lecture 18 of Spatial Statistics and Spatial Econometrics. We are going to take the lead from the previous lecture, where we relaxed one of the assumptions of a classical linear regression model.

In the previous lecture, we relaxed the assumption that the model errors are homoscedastic. And we introduced heteroscedasticity into the variance-covariance structure of model errors through the spatial dependence on prices of homes for example or groundwater levels etcetera, right?

In today's lecture, we are going to sort of, we are going to take another step and we are going to relax assumption A 2 which in lecture 16 we established as an assumption fundamental for causal inference.

So, to be able to move from correlation to causation, we require that the expectation, the conditional expectation of model errors is 0. That is expectation u_i given x_i is 0, where x_i is an explanatory variable, alright. And we also saw that that assumption is also important for the unbiasedness of a least squares estimator.

So, in the previous lecture, what we studied was a generalized least squares, estimator. Remember, that had got nothing to do with unbiasedness or causality. It was to reconcile the heteroscedastic model error structure.

$y_{i,j} \equiv y(i,j)$

$Spatially - lagged\ variables\ in\ a\ regular\ lattice$

$Shift\ Up \quad y_{i-1,j}$

$Shift\ Down \quad y_{i+1,j}$

$Shift\ Left \quad y_{i,j-1}$

$Shift\ Up \qquad y_{i-1,j}$

$Shift\ Right \qquad y_{i,j+1}$

$However,\ there\ is\ no\ analog\ for\ the\ irregular\ spaces\ (like\ for\ graoundwater\ level\ data).$

$This\ is\ why\ we\ study\ the\ notion\ of\ spatially\ lagged\ variables.$

So, before we sort of move forward and directly deal with the problem of causal inference, let us get back to specifying spatial dependence with the notion of spatial lags.

So, in the previous lecture, we had that is lecture 17, we had said that we could imagine a classical regression model, a traditional regression model that we are used to where each value of dependent variable y_i has an index i. This does not only provide us an ID for whom this value y is measured but also the location.

So, in that case, we could say you know y i, j which is equivalent to saying y at the coordinates i comma j. So, this you know enrichment of the notation is explicitly to account for spatiality in data.

And then if you have a regular lattice? What is a regular lattice? Well, a regular lattice is where you have equal, you know you have equal-sized cells constructed by n rows and m columns in general. That is a regular lattice.

That is to say that if I am considering the value of y_i at givens any given cell, I will be able to find its neighbor in the north, one step up equal size cell. So, you know if I take a step down, that is southward, I will find another neighbor. Similarly, a neighbor on the east and a neighbor on the west, right?

So, this idea can be notionally represented as what you have already understood. Shifting up would mean y_i minus 1 and j, right? Then, you have shifting down we are calling it y_i plus 1 j. So, i is representing the vertical direction, downward south is conventionally taken to be positive. It is you know; it is defined that way you could have also had it and vice versa.

Similarly, shifting left will give me the same i, but it will bring my j down from j to j minus 1, and on the right, I will have y_i, j plus 1, right? So, all of you are conversant with this by now.

Now, the issue is that this is a very stylized type of spatial structure. The data do not always exist in such nice niceties so far as the spatial structure is concerned. A very good example or case in point is the groundwater monitoring data that we have seen throughout this course, right?

So, we have seen that data for the state of Uttar Pradesh. So, if you, if you pay attention like I am going to draw an approximate shape of you, know of that state and you know it looks like something like as following. It is not to scale, not to shape, it is just an approximation, right?

And there when we look at the groundwater data, we saw a very dense cluster of wells on the west and when we move to the south you have a generally sorry on the east, we have a generally sparse model sorry monitoring network. And the monitoring network becomes denser around urban areas, right? So, they are just our observations till now.

So, in that case, we do not have a nice representation of a regular lattice in the case of real-world data. In that case, if you want to specify spatial dependence, we bring in the notion of spatially lagged variables. So, we bring in the notion of spatially lagged variables.

$In\ order\ to\ construct\ spatially\ lagged\ variables\ we\ formalize\ the\ local\ similarity$

$for\ describing\ the\ structure\ of\ interaction\ among\ the\ spatial\ units.$

$The\ quantitative\ unit\ that\ formalizes\ interactions\ with\ neighbors\ at\ any\ given\ location$

$in\ space\ is\ called\ as\ the\ Spatial\ Lag.\ For\ example.$

$$G_{i,L} = w_{i,1}G_1 + \ldots + w_{i,n}G_n = \sum_{j=1}^{n} w_{i,n}G_j$$

$Location\ is\ not\ her\ own\ neighbor\ i.e.,\ w_{ii} = 0 \qquad \forall\ i \epsilon \{1, 2, \ldots, n\}$

$Constructing\ "Spatial\ Lag"\ requires\ domain\ knowledge\ but\ it\ sumarizes\ pairwise$

$interactions,\ \ which\ would\ otherwise\ remain\ unidentified.$

$Potentially,\ n\, x \frac{(n-1)}{2} pairwise\ ineractions\ but\ only\ n\ groundwater\ observation$

$locations.$

$Matrix\ notation$: $\quad G_L = WG \rightarrow SPATIAL\ WEIGHTS\ MATRIX$

So, let us go ahead and look at an example where we will try to construct these spatially lag variables for groundwater data.

To construct spatially lagged variables, we formalize the local similarity for describing the structure of interaction among the spatial units. So, that is a sentence that is saying quite many things. So, let us break it down.

So, we have, we are talking about the structure of interaction among spatial units. So, what are we trying to do? We are trying to describe a structure of interaction. And we are doing that through this idea of local similarity. So, the fact that we believe in local similarity or local stationarity, you know we are designating spatial neighbors as the ones which will also provide a spillover effect on the value that is being measured.

So, the local similarity is in terms of the value observed, so the value observation. Value observed or you know the groundwater level observed, the price observed, or whatever, right?

We are describing it through the structure of interaction among spatial units. So, we are saying that units that are they are nearby might have similar values. This is also the idea of local stationarity, right? So, we are when we are constructing these spatially lag variables, we are following up on the ideas that we have developed till now, right? So, we are now formalizing it.

So, the quantitative unit that formalizes interaction with neighbors at any given location in space is called spatial lag. Very very important. So, we are now introducing an entity called spatial lag. It is an entity that formalizes interactions with neighbors for any given location in space.

For example, I am looking at, I am defining a spatially lag groundwater level at location i. So, I am going on to location I, and I am searching for its neighbors. So, some wells are located near i, and some others are located farther away from i. But I am interested in the spatially lagged value at location i.

Remember the groundwater level that is observed at location i is given as G of i. We are trying to create a lag. So, we are somehow summarizing information about what is happening around this location i. So, G_{iL}, where L is a representation of the spatial lag is equal to w_i, 1 G_1 plus all the way to win G_n which is equal to summation j equals 1 win G_j, right?

I am able to provide weight to each neighbor. So, in case if for example, we have this neighbor k which is far away, and we do not believe that this is going to have any spatial dependence of groundwater level at k on what is happening at location i, then you know I can conveniently say w_{ik} is 0. But that you know specifying it to be 0 or nonzero is another matter.

The point of point that I am trying to make here is that this formulation of creating or constructing spatial lags with the help of these spatial weights, you know allows me to generalize the idea of neighborhood spillovers or neighborhood effects, right?

I do not have to go in and specifically figure out who has how many neighbors. I can at least mathematically or notational provide a general formulation of what a spatial lag means. So, spatial lag is nothing, but values in the neighborhood, values in the neighborhood, such that they are weighted by a factor w i, j for the j^{th} neighbor or the j unit which is a neighbor of i in the given domain, right?

In this formulation, everybody is designated a neighbor, right? A unit that is approximate to i is also considered a unit, and which is farther away from i is also considered a neighbor. What differentiates these two units which are approximate or farther away is this weight w i, j. And so, this nice you know representation j equals 1 to n w i, j sorry for the typo w i, j G_j provides me a generalized formalization of the spatial lag effect.

So, if I am trying to understand what is the total extent of spatial effects on the groundwater level at location i, I should study this spatial lag you know entity summation j equals 1 to n w i, j G_j, right? Concisely we write as G_i capital L. This capital L is just a lag representation.

One thing to note here is that a location, any location is not set own neighbor. So, location i is not her own neighbor; that means, w_{ii} will always be 0 for all i. So, w_{ii} will be 0 for all i.

Even though w ii is 0 my general formulation does not change. I can always for any i be it i equals 1, 2, 3, 4, 5, n, I can use this formulation. The only matter to understand here is if I am talking about the first location, location 1, location ID 1, then w_{11} will be 0. If I am talking about location ID 2, then w_{22} will be 0, and so on and so forth.

So, overall constructing these spatial lags requires domain knowledge, right? But it summarizes pairwise interactions which would otherwise remain unidentified. So, the spatial lag is characterizing pair-wise interaction. So, i and j in a pairwise sense the way they interact is summarized by the spatial lag. It is a component of the construction of the spatial lag variable, right?

But constructing it requires domain knowledge. How and why? Well, what value should w i, j carry? Should it be just 0 and 1? Well, should it be the nearer value should get larger, you know a value of weights than the ones that are farther away? And how much? By how much? What is the difference between different weights? Just because two points are equidistant to a location i, will they have the same spillover effect?

Well, these are matters of complexities and that is why if you are working with groundwater data, you know just like we have seen in the case of stationarity we require a lot of very strong domain knowledge. Here too we require strong domain knowledge, right?

So, if you are working with groundwater data, it's best to work with a hydrologist. If you are working with agricultural data, it is probably best to consult with agronomists and so on and so forth.

If you are working with geological data, like cold exploration data or oil exploration data is probably best to consult with geologists. If you are working with population density data, it is probably best to consult with population scientists, right? So, whatever data set whichever domain your data set belongs to, one should read up on the literature in that domain to do justice in defining these weights.

Now, potentially you have n times n minus 1 by 2 pair-wise interactions, but only n groundwater observation location. So, you have a lot of pair-wise interactions that you have to worry about, right?

So, for every i, there are n potential neighbors. Of course, herself is not a neighbor. So, you can remove it, but in this formulation, I am counting herself also a neighbor just because of the general form that I am studying. Just because w_{ii} is 0, I am not keeping it out. I am just saying you know because it is not her own neighbor, i is not her own neighbor, I can simply assign or define w_{ii} to be 0, but then keep my formulation of G_{iL} to be the same.

But potentially I have n times, n minus 1 by 2 which is much greater than n pairwise interactions to study, right? So, if you have so many degrees of freedom, and probably so many unknowns, so many parameters to estimate, then probably it is not going to be very efficient.

I am also providing a matrix notation of the same entity here. So, I am calling G_L a general lag spatial lag matrix. I am not calling it G_{iL} anymore. I am calling it G_L is an n by 1 matrix, so we can write this down, for our understanding.

$$G_L = WG$$

$$\left[G_{1L}\; G_{2L}\; \vdots\; G_{iL}\; \vdots\; G_{nL}\;\right] = \left[w_{11}\; w_{12}\; \dots\; w_{21}\; w_{21}\; \dots\; \vdots\; w_{i1}\; \vdots\; w_{n1}\; \vdots\; w_{i2}\; \vdots\; w_{n2}\; \vdots\; \dots\; \vdots\; \dots\; w_{1i}\right.$$

$$G_{iL} = \sum_{j=1}^{n} w_{ij} G_j \rightarrow i \in \{1, 2, \dots, n\}$$

So, I am writing G_L equals W times G. G_L is considered to be n by 1, right? W is n by n and G is n by 1. So, W is n by n and G is n by 1, right? So, I have a lag matrix. So, I have G_{1L}, 2L, and keep going G_{iL}, right? G_{1L}, G_{2L}, keep going till G_{IL} then I have finally, G_{nL}.

This is an n by 1 matrix. This is equal to an n by n matrix of weights, so each row will correspond to the location 1, 2, i, n. On the columns, I am looking for neighbors. So, everybody is a neighbor to each other. The only thing is we have a weights understanding of them, right? So, I have 1, 2, and keep going i, all the way to n.

So, columns are potential neighbors and rows are the entities for which we are trying to search neighbors for. This is again n by n and then G is just the location, the column of values observed at each location in my data set.

So, now w_1 and its neighbor w_{11} which is going to be 0, we have learned that. Then, you have w_{12}, w_{1i}, all the way to w_{1n}. So, I have weights, like a row of weights sitting there as you know w_{11}. Then, I have for 2, w_{21}, w_{22}, w_{2i}, all the way to w_{2n}. Similarly, w_{i1}, w_{i2}, w_{ii}, w_{in}. Then, w_{n1}, w_{n2}, w_{ni}, w_n.

So, now I am multiplying n by n matrix by n by 1 matrix. As we have seen earlier, they are conformable and perfect. The overall product will be n by sorry about that. So, they are conformable because the number of columns on the W matrix is equal to the number of rows on the G matrix.

And, If I multiply them, I am going to get an n by 1 matrix, right? Something we are very well aware of. So, let us write it down. So, I have an n by 1 that I am looking for. So, the first column is to be multiplied by the first row to find the first cell of this n by 1 matrix.

So, this is going to be w_{11} G_1, w_{12} G_2, w_{1i} G_i plus w_{1n} G_n. Similarly, the second one will be summation i equals or j equals 1 to n, w_{2j} G_j, right? So, I am simply going to multiply the second row with the first column to get the second row of this n by 1 matrix. Similarly, I am going to have j equals 1 to n w i, j G_i, and then finally, the last element will be j equals 1 to n w_{nj} G_j, right?

So, now I can see that G_{iL} equals, the summation G_{iL} equals summation j equals 1 to n w i, j G_j which is the same as the scalar formation that we had seen on the previous page here.

So, by now I am sure that you are very conversant between scalars and matrices, scalar forms and matrix forms of expressing the same linear equations. But you know it is very critical to keep learning and keep improving our translation between the two devices. That is why I decided to sort of you know solve this for you.

My request is that you please go over this on your own time. So, that it becomes absolutely clear you know for your purposes, right?

As we go forward we are going to move to the matrix of formulation more and more because the matrix is so concise, I mean you have this type of a bulky G_{iL} equals summation j equals 1 to n w i, j G_j, such a bulky notation in scalar form for all I's going from 1 to n. All of these n equations; so, you have n equations sitting here are summarized in this simplistic matrix form, right?

So, matrix form, matrices are very useful, and very handy when we are dealing with bulky notations. So, it is very important to learn to translate between scalar and matrix notations. So, I request you to write it out in your own time, so that it becomes more and more natural going forward.

$$W = [0\,1\,0\,1\,0\,0\,0\,1\,1\,0\;\;0\,1\,1\,0\;\;0\,0\,1\,1\;\;\;\;1\,1\,0\,1\,1\,0\,0\,0\,1\,0\;\;1\,1\,0\,0\;\;1\,0\,0\,0\;]_{6\,x\,6}$$

$Binary\ contiguity\ weights\ matrix$

$$w_{ij} = \{0\ if\ i\ and\ j\ are\ not\ neighbors\ 1\ if\ i\ and\ j\ are\ neighbors$$

So, the next step is to construct a spatial weights matrix. So, now we look at an example where there are some spatial units, units 1, 2, 3 4 5, and 6 are assembled in a given spatial structure.

What is the spatial structure? The spatial structure is such that 1 has is neighbors with 2, with 4, and with 5. So, 1 is a neighbor with 2, with 4, and 5. 2 is a neighbor with 1, with 5, and with 4, right?

So, 2 is also neighbors with 1, with 5, and with 4. 5 on the other hand is connected with 1, 2, 4, and 3. So, 5 has 4 neighbors unlike you know 1 and 2, 5 has many more neighbors. So, 5 has you know 1, 4, 2, and 3 as neighbors. 3 on the other hand has only two neighbors, 5 and 6. 6 finally, has just one neighbor that is 3, right?

Now, this is a complex spatial structure. There are entities distributed in space, right? Location is deterministic. It is not random. So, once you are located as a neighbor of 2, 4, and 5 that is what it is for you, right? So, if it is a home that is a large mansion at location 1 and it is connected to these 3 different homes home 2, home 4, and home 5, then their location are sort of fixed.

You can only sort of you know do as much to change the neighborhood of location 1, right? So, this is the spatial constraint that is driven by the links that are exogenously driven. We do not get to choose who is connected to whom. What we get to do is we get to summarize that in a spatial weights matrix.

So, what does this spatial weights matrix do? First of all, you can see the diagonal elements are 0, these are $w_{ii's}$ which are 0. So, home 1 or location 1 is not her own neighbor, location 2 is not her own neighbor, location 3 is not her own neighbor, and so on, till location 6 not being her own neighbor.

If we have 6 different entities the size of the weights matrix is 6 by 6, we are working with a 6 by 6 matrix, right? Whenever we have a neighbor we provide a weight of 1 and if they are not neighbors we provide a weight of 0.

So, if I go back to location 1, I figured that 1 has, 1 is a neighbor with 2, 4, and 5. So, at locations 2, 4, and 5 in the columns, I have 1s and for the rest, I have 0s at 3 and 6, right?

Now, if I look at it the number of 1s that I find in rows, the number of 1s provides me an understanding of how networked or how richly networked each location is. At an extreme is location 6 which only has 1 neighbor which is number 3, right? So, 6 only has 1 neighbor, the rest all are non-neighbors including herself, right? So, that is how a weights matrix is able to concisely account for the neighborhood structure.

$SPATIAL\ WEIGHTS\ TRANFORMATIONS\ -\ ROW\ STANDARDIZED\ WEIGHTS$

$Weights\ are\ typically\ rescaled\ such\ that\ \sum_j w_{ij} = 1$

$$\tilde{w}_{ij} = \frac{w_{ij}}{\sum_j w_{ij}}$$

$W = [0\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 1\ 0\ \ 0\ 1\ 1\ 0\ \ 0\ 0\ 1\ 1\ \ \ 1\ 1\ 0\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ \ 1\ 1\ 0\ 0\ \ 1\ 0\ 0\ 0\ ]\ \ \tilde{W} = [0\ 1/3\ 0\ 1/3$

Now, we do not use these weights matrices. We define them as the way we have seen them till now. We assign an entity 1, if j is my neighbor, and 0, if j is not my neighbor, for each row, for each i in my data I am going to do that.

But I then row standardize these weights. What does it mean to row standardize? It means that each row must sum to 1. So, if I have a row with 3 1s and the rest are 0s, I am going to multiply each of these by the sum of what is found on that row. So, each value of 1 is now normalized by 3. So, it becomes one-third, one-third, one-third. Of course, the 0 is also divided by 3, but 0 divided by 3 is just a 0.

If I look at number 5 it has 4 neighbors, right? Now, what I do is I sum the row I get 4, and I divide the whole thing by 4. So, I get 1 4, 1 by 4th, one-fourth, one-fourth, 0, and 0 in that row, in row number 5.

Similarly, if I go and look at row number 2, again I have 3 neighbors, overall for entity 2, I have only 2 neighbors. For Entity 3, so I have one-half and one-half. I have 3 neighbors for Entity 4 and I have only 1 neighbor for Entity 6. So, I have a value of 1, which is 1 divided by 1, right? So, the W prime is known as the rho standardized weights matrix. It is called the row standardized weights matrix.

$$G = \rho_g WG + X\beta + \rho_x WX + u$$

or

$$G_i = \rho_g \sum_{j=1}^{n} w_{i,n} G_j ++ X_i\beta + \rho_x \sum_{j=1}^{n} w_{i,n} G_j + u_i$$

A question arises why do we row standardize these matrices? Well, to understand that let us look at the following model.

This model is a spatial regression model. What is happening in this model is that I am multiplying, I am modeling G which is an n by 1 vector as a function of a lag of G, right? So, this is G_L which is nothing, but again an n by 1 matrix, where W is n by n and G is n by 1, right?

So, I am regressing G on G_L, X which are some exogenous variables, and then also WX. So, I am all not only spatially weighing the G values which is the dependent variable, but also the experimental variable. This is a very general structure that I have put in front of you.

Now, this is going to be n by k. So, I have an n by k sitting here, right? And then I have the model error u, which is going to be n by 1. The parameters rho G and rho X are called spatial coefficients, or coefficient parameters. These parameters explicitly account for the spatial spillovers, and beta is my model parameter as I am already aware of. It is a coefficient sitting on each X value, right? u is just a model error. Everything that I did not account for in this model will go into this u, u is going to be a random number.

The question that we are asking is, why do we require W to be row standardized? Well, what happens is that you know I am trying to understand the impact of the lags, the spatial lag on the value at location i. So, I want to know how much groundwater levels in the neighborhood impact groundwater level values at location i.

In this, quest, you see in this formulation this impact will be generated from two pieces of information. One is W and the second is G, right? What I am really interested in is this G and not in W, that is to say, my quest as an analyst is not about understanding whether more neighbors or higher connectivity causes groundwater levels to increase or decrease, but to understanding what happens with the levels in the neighborhood on the levels at location i.

Because I want to connect groundwater levels with groundwater levels in a neighborhood, I want to normalize the effect of the number of neighbors. I do not want the effect to be driven by the number of neighbors.

In case I do not use you know, I do not standardize $w_{i's}$, then what happens is that just because some entities have 4 neighbors this w_i value sums us 4 times, right? And just because some other entity has just 1 neighbor which is entity 6 in our previous example, they will have a smaller spillover, just by the virtue of low connectivity.

Well, that is not what I am trying to learn here. I am trying to learn the linkage between groundwater levels in a neighborhood on a location i, right? So, to filter out the effects due to the degree of connectivity, I row standardize my spatial weights matrix. So, I hope that is clear.

$Note\ that\ each\ location\ i\ is\ not\ her\ own\ neighbor,\ i.e.,\ w_{ii} = 0$

Now, a property of this spatial lag is that it is similar to, but not the same as window average in the case of time series data. So, I have an example of a time series of sales per quarter.

Now, the time series is just joining all the different realizations of sales in different quarters starting from 1986 to 1996 for some entity, right?

The window average says go to quarter 3 of your quarter 4 of 1992 till quarter one of 1994, right basically takes 6 quarters and sums everything between them. So, this scanner will move step by step, this scanner will move one step here and it will look like the following at you know at t. So, this will be for t plus 1.

So, now, this scanner just averages everything between these windows. And this window, this moving average window is moving as we go along the time periods.

Now, spatial lag if you think about it is doing something similar. It is taking a weights type of a spanner or a scanner side window average type of a like a window which goes through every value whenever I am standing at a value i and it just takes average, which is a weighted sum, which is weighted average.

The only difference is that here the value at the middle itself which is here in the case of this time t and t plus 1, these things are not counted in the weighted average, right?

So, the value at the location at which we are conducting this weighted average exercise or this window average exercise, that value itself is excluded in case of the spatial lag construction, right? So, the fact that each location i is not her own neighbor means that the spatial lag is not a window average. It excludes the middle value. It excludes the value at the location where I am standing wherein I am conducting a local average.

$$G = \rho W G + X\underline{\beta} + \underline{u}$$

$$\underline{G} = X\underline{\beta} + \rho_x W X + \underline{u}$$

1.Spatially lagged dependent variables (W.G)

- Spatial Lag (or autoregressive) Model

2.Spatially lagged explanatory variables (W.X)

- Spatial cross-regressive model or SLX model

3.Spatially lagged error terms ($W.e$)

- Spatial Error Model

$\{\underline{G} = X\,\underline{\beta} + \underline{u}\, s.t: \quad \underline{u} = \rho w\underline{u} + \underline{\varepsilon}$

So, a little bit more into, you know just a little introduction before we will go into these things. I just want to introduce some of these variables that are constructed for a regression model.

First is called a spatially lag-dependent variable. Now, G being a function of rho WG, this here is called a spatially lag-dependent variable. And this, when it is added to the model that we are used to which is you know X beta plus u if I were to remove rho WG_i will get a model that I am used to.

When I include a spatial lag-dependent variable such a model is called a spatial lag model. It is a specific definition, as it is a spatial lag model. Instead of rho WG, if I were to include X beta plus rho X, W X plus u, then such a model includes spatially lagged explanatory variables. So, now, I have lagged behind the explanatory variables, so I am saying the spatial effects are coming from the explanatory variables.

So, in the case of house price data, the spatial effects are not coming directly from the prices of the homes in the neighborhood, but from the number of rooms in homes in the neighborhood or the public amenities for the homes in the neighborhood, right? So, I am sort of you know going over the example that we covered in the last class extensively, right? In that case, we call such a model the slx model or the spatial cross-regressive model.

And if instead we were to include these effects in the error structure such that u is rho W u plus epsilon then such a model is called a spatial error model. Going forward we will go over each of these models one by one. We will look at what they mean, what is the consequence of including or not including these effects, how we choose between these models and so on and so forth.

We will conduct that exercise. But for now, I just want because we have introduced spatial weights I wanted to you know also introduce these 3 model variants of spatial regression models.

So, that is it for this part of the lecture. In the next part of lecture 18, we will be looking at what we started with as the departure from the classical assumption of causality that ensures causality which is assumption 2 of a linear regression model. And we will see how we obtain causal inference in spatial regression models.

So, thank you very much. And see you in the next part.