Description Remarks and examples Reference Also see

Description

Spatial lags and spatially autoregressive errors are defined by the spatial weighting matrix **W**. This entry describes the weighting matrix.

You may also be interested in introductions to other aspects of Sp. Below, we provide links to those other introductions.

Intro 1	A brief introduction to SAR models
Intro 3	Preparing data for analysis
Intro 4	Preparing data: Data with shapefiles
Intro 5	Preparing data: Data containing locations (no shapefiles)
Intro 6	Preparing data: Data without shapefiles or locations
Intro 7	Example from start to finish
Intro 8	The Sp estimation commands

Remarks and examples

Remarks are presented under the following headings:

Understanding the W matrix
Missing values, dropped observations, and the W matrix

Understanding the W matrix

You will usually construct **W** on the basis of shapefiles (maps) that you obtain over the web or from other sources. It is so easy to do that you might think you can ignore the details of **W**. You cannot. You need to understand **W** to interpret results from the models you fit. Moreover, those models are conditioned on **W**, and the matrices you use are as much a part of your model as are the variables you include or intentionally exclude.

You use **W** in your models in three ways:

1. You include $\lambda\mathbf{Wy}$ to allow nearby outcomes to affect outcomes.
2. You include $\gamma\mathbf{Wx}$ to allow nearby covariates to affect outcomes.
3. You include autoregressive errors $(\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\epsilon}$ to allow nearby errors to affect outcomes.

You can think of **W** as specifying the potential spillover as long as you realize that the actual spillovers are as follows:

1. The effect that y_i of area i has on nearby y's from the term $\lambda\mathbf{Wy}$.
2. The effect that x_i has on nearby y's, both from the term $\gamma\mathbf{Wx}$ and from the effect that x_i has on y_i, which in turn affects nearby y's.
3. The effect of including an autoregressive error.

The weighting matrix **W** is effectively a constraint placed on the individual spillovers formulated as part of the model specification.

For instance, if $W_{1,3}$ is 0, then there will be no spillover from 3 to 1 contributing to the total. It is constrained to be 0. If $W_{2,6}$ and $W_{4,7}$ are both 1, then individual spillovers from 6 to 2 and from 7 to 4 will be constrained to be equal. If $W_{5,7}$ is 2, then the spillover from 7 to 5 will be twice that of 7 to 4.

To see how this works, we will consider the matrix for four fictional places:

- Mordor, a dark land in J. R. R. Tolkien's *The Lord of the Rings*.
- Bree, a village from the same story.
- Hogsmead, a village from J. K. Rowling's *Harry Potter* novels.
- Hogwarts, a school near Hogsmead in the *Harry Potter* novels.

Spatial weighting matrices have 0s down the diagonal:

Spatial weighting **W**

	Mordor	Bree	Hogsmead	Hogwarts
Mordor	0			
Bree		0		
Hogsmead			0	
Hogwarts				0

The 0s down the diagonal may surprise you. Perhaps you expected 1s. $W_{i,j}$ is the spillover from j to i, so $W_{i,i}$ is the spillover from i onto itself. Surely, geographic area i affects itself. Your thinking is correct, but you forgot that the purpose of **W** is to specify the effect of nearby areas. You will measure the effects of i on itself by adding other variables, such as **x**, to your model:

$$\mathbf{y} = \beta_0 + \beta_1\mathbf{x} + \beta_2\mathbf{Wx} + \cdots$$

In this model, β_1 measures the effect of x_i on y_i, and $\beta_2\mathbf{W}$ measures the effect of $x_{i'}$ from other areas $i' \neq i$ on y_i. **W** has 0s down the diagonal so that **W** serves its intended purpose.

A **W** matrix could contain all 0s:

Spatial weighting **W**

	Mordor	Bree	Hogsmead	Hogwarts
Mordor	0	0	0	0
Bree	0	0	0	0
Hogsmead	0	0	0	0
Hogwarts	0	0	0	0

If the matrix contains all 0s, there are no spatial effects. The observations are independent, and you may as well use `regress` to fit the model.

You use Sp estimation commands when some elements of **W** are nonzero. Zeros are nonetheless a reasonable value for many of the elements. For instance, Mordor and Bree are from one set of novels, while Hogsmead and Hogwarts are from another. It would be reasonable to assume (to constrain) that there are no spillover effects between them. We would have the following matrix:

Spatial weighting **W**

	Mordor	Bree	Hogsmead	Hogwarts
Mordor	0	?	0	0
Bree	?	0	0	0
Hogsmead	0	0	0	?
Hogwarts	0	0	?	0

In the above matrix, we are specifying that Mordor and Bree are independent of Hogsmead and Hogwarts, and vice versa. The question marks stand in for the values left to be filled in, which are

- $W_{1,2}$, the potential spillover of Bree on Mordor.
- $W_{2,1}$, the potential spillover of Mordor on Bree.
- $W_{3,4}$, the potential spillover of Hogwarts on Hogsmead.
- $W_{4,3}$, the potential spillover of Hogsmead on Hogwarts.

Nonzero values in **W** must be positive. The larger the value in W_{i_1,i_2}, the more the potential spillover.

How shall we measure spillover? It turns out not to matter so long as we are consistent. Said differently, only ratios of elements in the matrix matter. Remember how spatial lags are used:

$$\mathbf{y} = \beta_0 + \beta_1\mathbf{x} + \beta_2\mathbf{W}\mathbf{x} + \cdots$$

Fitted coefficient β_2 measures the effect of the spatial lag. If we replaced **W** with $2\mathbf{W}$, the result would be to halve β_2, just as β_1 would halve if we doubled **x**.

We will set $W_{3,4}$, the potential spillover of Hogwarts on Hogsmead, to 1, and in setting this first nonzero value, we have decided on the units. The units are Hogwarts on Hogsmead. If we set an element to 2, then we are setting the potential spillover to be twice that of Hogwarts on Hogsmead. If we set an element to $1/2$, then we are setting the potential spillover to be half that of Hogwarts on Hogsmead.

If we also set $W_{4,3} = 1$, we will be constraining the potential spillover of Hogsmead on Hogwarts to be the same as Hogwarts on Hogsmead. Our matrix would be

Spatial weighting **W**

	Mordor	Bree	Hogsmead	Hogwarts
Mordor	0	?	0	0
Bree	?	0	0	0
Hogsmead	0	0	0	1
Hogwarts	0	0	1	0

What should we make the spillovers of Bree on Mordor and of Mordor on Bree? In the *Lord of the Rings* story, Mordor is far from Bree, larger than Bree, and actively exporting evil at the speed of magic. Bree, meanwhile, is a speck that Mordor could brush away with little effort.

We will set $W_{2,1}$, the spillover of Mordor on Bree, to 4 and $W_{1,2}$, the spillover of Bree on Mordor, to 0.1. The spillover of Mordor on Bree will be four times that of Hogwarts on Hogsmead. Meanwhile, the spillover of Bree on Mordor will be one-tenth that of Hogwarts on Hogsmead.

You might well question the numbers we have chosen. Why is the spillover of Mordor on Bree 4 and not 5? Or 10? We have no satisfactory answer, and that is why in real problems researchers often set potential spillovers to 1 for adjacent areas and to 0 elsewhere, or set potential spillovers to the inverse of the distance between the locations. Both seem more defensible, although defending them can be problematic. Do second-order neighbors really have no effect? Or in the case of inverse distance, why not inverse distance squared? Sometimes theory can provide an answer. The spillover of a light bulb is inverse distance squared. In other cases, there are no satisfactory answers except that making partially justified assumptions and accounting for spillover effects is preferable to assuming that spillover effects are all 0.

So we have

Spatial weighting **W**

	Mordor	Bree	Hogsmead	Hogwarts
Mordor	0	0.1	0	0
Bree	4	0	0	0
Hogsmead	0	0	0	1
Hogwarts	0	0	1	0

And the final **W** matrix is

$$\mathbf{W} = \begin{bmatrix} 0 & 0.1 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

We could enter this matrix into Sp and proceed to estimation. It would be easy enough to do; see [SP] **spmatrix userdefined**, [SP] **spmatrix fromdata**, and [SP] **spmatrix spfrommata**.

Although we could do this, in reality you will not. You will have more than four geographical units—you might have 3,000 counties. To say nothing of the misery of entering a 3000 × 3000 matrix, you are not going to carefully consider and research all 8,997,000 pairs of counties. You are going to assume that only adjacent counties affect each other—called "contiguity" in the literature—or that spillover effects are proportional to the inverse of distance between counties. You are going to do that because you can create such **W** matrices by typing a single command such as

```
. spmatrix create contiguity Wc          // contiguity
. spmatrix create idistance  Widist      // inverse distance
```

We told you that the units in which the weights are measured do not matter, but that is not exactly true. They do not matter if you only include spatially lagged covariates. If, however, you use the spatial weighting matrix to lag the dependent variable ($\lambda\mathbf{Wy}$) or for autoregressive errors [$(\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\epsilon}$], then $\widehat{\lambda}$ and $\widehat{\rho}$ will be easier to interpret if the matrix is scaled appropriately. In that case, $\widehat{\lambda}$ and $\widehat{\rho}$ should be between -1 and 1 unless the solution is explosive.

Explosive solutions can arise in spatial analysis for the same reasons they arise in time-series analysis. If A affects B and B affects A, and if the coefficients are large enough, then feedback becomes amplified. A sends a large value to B, which B receives, amplifies, and sends back to A, whereupon the procedure repeats, and eventually, the process explodes in a mess of infinities.

To fit models, the Sp software virtually requires that you scale the matrices used to produce lags of the dependent variable or autoregressive errors, and you ought to scale the other matrices too. The software produces more accurate results when inputs are scaled.

Scaling is so important that when you type the commands

```
. spmatrix create contiguity Wc
. spmatrix create idistance Widist
```

scaling is performed automatically, and you have to go out of your way to prevent it, which you can do by typing

```
. spmatrix create contiguity Wc, normalize(none)
. spmatrix create idistance Widist, normalize(none)
```

By default, weighting matrices are scaled so that their largest eigenvalue is 1. See [SP] **spmatrix create** and *Choosing weighting matrices and their normalization* in [SP] **spregress** for details about normalization.

Missing values, dropped observations, and the W matrix

Missing values sometimes appear in data. When fitting models with such data, the usual solution is to omit the observations from the estimation sample. That can be justified when observations are independent, but observations are not independent in SAR models.

Spatial models allow for spillover effects from nearby areas. In spatial data, observations are areas. Omitting some areas means that the spillovers from them are no longer being included in the fitted model. Consider two adjacent counties and assume that one of them is dropped from the estimation. Then, the spillover from the dropped county to its neighbor—a neighbor still in the data—becomes 0 even though there really is spillover. It is just unobserved spillover.

Thus, Sp estimation commands handle missing observations differently from Stata's other estimation commands. If an area is defined in a spatial weighting matrix and that area is not observed in the data, Sp refuses to fit the model unless you specify option `force`.

Imagine that you type

```
. spregress y x, gs2sls ivarlag(W: x)
```

and that observation 4 contains a missing value for `x`. Most Stata estimation commands would omit the observation from the estimation sample and proceed with estimation. `spregress` will issue an error and mention the `force` option.

```
. spregress y x, gs2sls ivarlag(W: x)
  (1412 observations)
  (1 observation excluded due to missing values)
  (1411 observations (places) used)
  (weighting matrix defines 1412 places)
weighting matrix defines places not in estimation sample
    Excluding observations excludes the spillovers from those observations to
    other observations which are not excluded. You must determine whether this
    is appropriate in this case and, if it is, specify option force.
r(459);
```

You would be on firm theoretical ground to specify the `force` option if the fourth column of `W` contained only 0 values, because in that case there are no spillovers from observation 4 to the other areas. Meanwhile, the fourth row does not have to be all 0s. The other areas might spillover to observation 4, but that will not bias results if observation 4 is omitted. It is the unobserved spillovers from observation 4 that cause bias.

You would be on muddy theoretical ground—which most applied researchers consider to be firm enough—if the fourth column of `W` contained only small values.

You would be sinking in a swamp if the fourth column of `W` contained any large values. We at StataCorp might go there, but if we did, we would afterward try replacing `x[4]` with various reasonable values to determine how sensitive our forced results would be to the missing observation.

Reference

Liu, D. 2017. How to create animated graphics to illustrate spatial spillover effects. *The Stata Blog: Not Elsewhere Classified.* https://blog.stata.com/2018/03/06/how-to-create-animated-graphics-to-illustrate-spatial-spillover-effects/.

Also see

[SP] **Intro 3** — Preparing data for analysis

[SP] **Intro 7** — Example from start to finish

[SP] **spmatrix** — Categorical guide to the spmatrix command

Stata, Stata Press, Mata, NetCourse, and NetCourseNow are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. StataNow is a trademark of StataCorp LLC. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2025 StataCorp LLC, College Station, TX, USA. All rights reserved.

For suggested citations, see the FAQ on citing Stata documentation.