**Homework-09-王慈俭-SA23008216**

**The questions:**

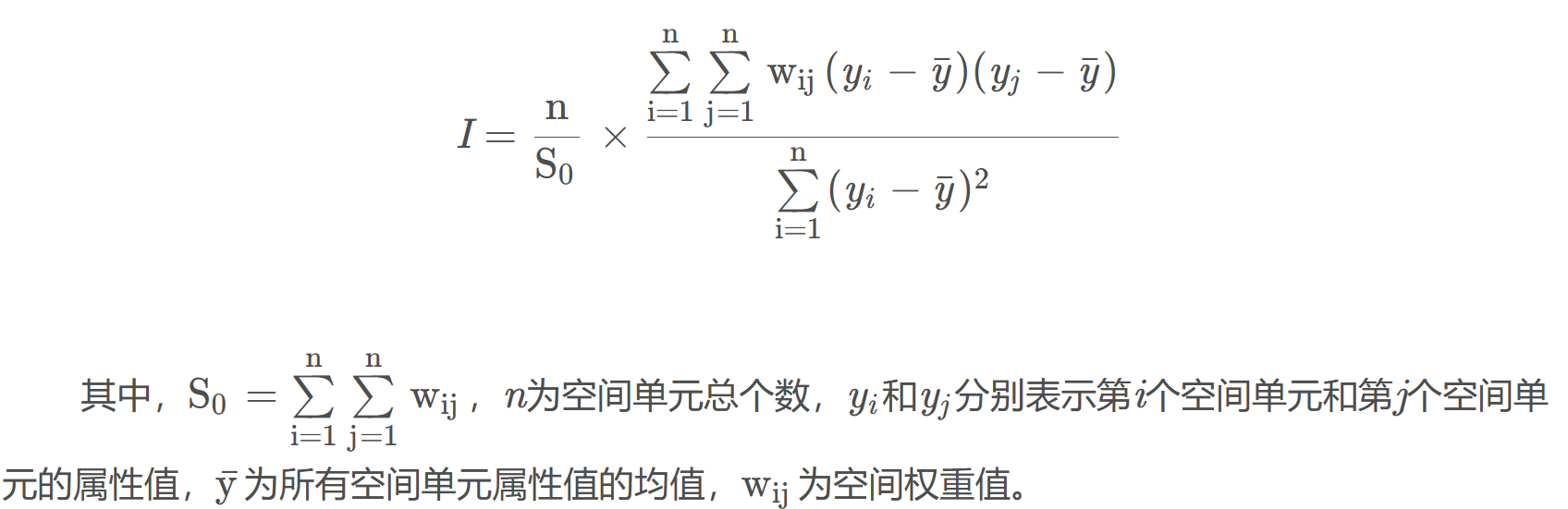
1. **Local estimates of Moran’s I and Geary’s C are often used to identify spatial clustering/hotspots) and outliers in attribute values. Please point out which conditions indicate hotspots and which ones indicate outliers.**

**当地估计的Moran's I和Geary's C经常用于识别属性值中的空间聚集/热点和异常值。请指出哪些条件表明热点，哪些表明异常值。**

答：数据的空间结构表现为空间相邻地点观测到的值之间存在相关性或不相关。在大多情况下，接近样点倾向于比距离更远的点有更相似的值，从而产生空间依赖性。因此，在不同样点观察到的群落物种组成差异要么是由群落内部过程产生的“空间自相关”，要么是由外部因素导致生物群落中的“空间依赖”。如果不考虑地理空间,就无法充分理解生物群落的内部结构及其对环境的反应。为检测空间模式，需先定义地点间的空间相邻关系，将这种关系存储在SWM(SWeighting Matrix)中,然后计算 Geary 指数(Geary's ratio)和 Moran 系数(Moran’s coefficient)，以度量空间自相关。

①Moran 系数

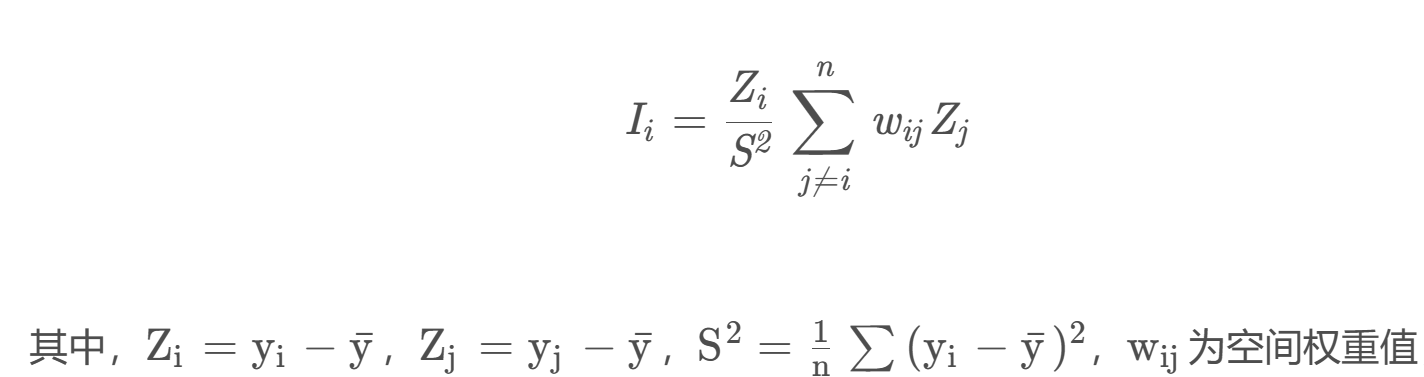
莫兰指数分为全局莫兰指数（Global Moran's I）和局部莫兰指数（Local Moran's I）。全局莫兰指数用来分析有没有空间自相关性存在，而局部莫兰指数用来探测异常值或者集聚出现的范围和位置。通常情况，先计算一个地区的全局莫兰指数，全局指数会告诉我们空间是否出现了集聚或异常值。如果全局有自相关出现，再做局部自相关，局部Moran's I会告诉我们哪里出现了异常值或者哪里出现了集聚。

Global Moran's I计算公式如下：

I的取值范围为[-1,1]，具体范围所代表含义如下表所示：

|  |  |
| --- | --- |
| I的范围 | 含义 |
| I>0 | 所有地区的属性值在空间上有正相关性，即属性值越大(小)越容易聚集在一起 |
| I=0 | 地区随机分布，无空间相关性 |
| I<0 | 所有地区的属性值在空间上有负相关性，即属性值越大(小)越不容易聚集在一起 |

在空间相关分析中，若Global Moran's I显著，则可认为该区域存在空间相关性。然后在通过计算Local Moran's I来探究具体哪些地方存在空间聚集（热点）/异常值。

Local Moran's I计算公式如下：

局部莫兰指数的取值范围在-1到1之间，具有以下含义：

Ⅰ正值表示该地理单元与其邻近地理单元之间存在正的空间自相关性，即该地理单元与其邻近地理单元的值相似程度较高，可能形成聚集模式。

Ⅱ负值表示该地理单元与其邻近地理单元之间存在负的空间自相关性，即该地理单元与其邻近地理单元的值相似程度较低，可能形成分散模式。

Ⅲ零值表示该地理单元与其邻近地理单元之间不存在明显的空间自相关性，即该地理单元的值与其邻近地理单元的值无关。

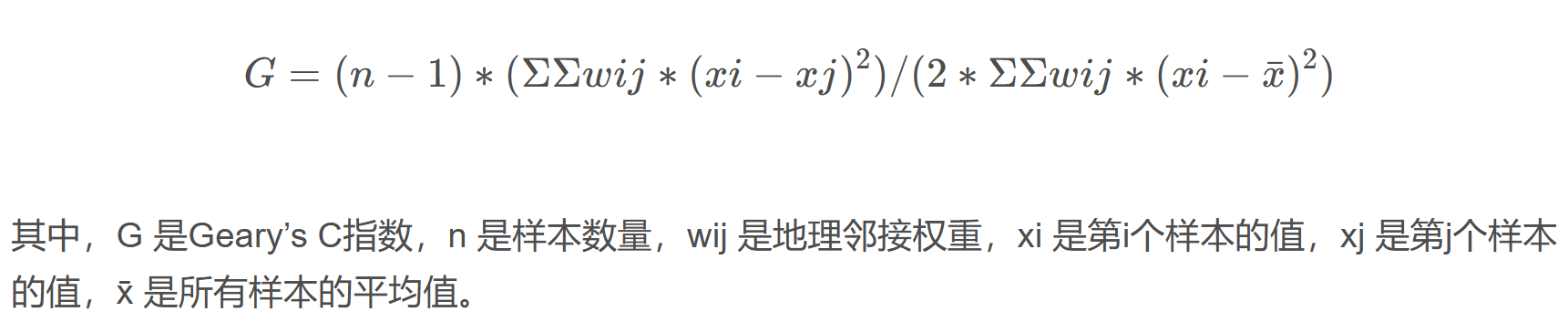
**热点区域（高正值）：区域的局部Moran's I值显著高于全局Moran's I值，并且周围区域的值也是相似的，则此区域为热点区域。这表示该区域内的观测值与周围观测值的值相似，存在空间聚集或热点。**

**异常值区域（高负值）：区域的局部Moran's I值显著低于全局Moran's I值，并且周围区域的值也是相似的，则此区域为异常值区域。这表示该区域内的观测值与周围观测值的值差异较大，存在空间离散或异常点。**

②Geary 指数

Geary’s C（也称为Geary’s coefficient）是一种用于衡量空间自相关性的统计指标，它可以用来评估地理数据中的空间聚集或离散程度。

Geary’s指数的计算公式如下：

Geary’s C指数的取值范围是0到2之间。当G接近于1时，表示数据呈现随机分布；当G接近于0时，表示数据呈现正相关的空间聚集；当G接近于2时，表示数据呈现负相关的空间离散。

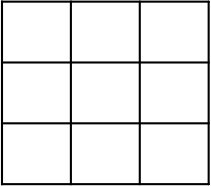
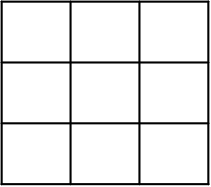
Geary’s C指数的计算需要考虑地理邻接权重，这是用来表示地理空间上的相邻关系的权重矩阵。具体的权重计算方法可以根据具体的研究问题和数据特点进行选择和定义。

**聚集/热点区域：当Geary’s C指数的值接近于0时，表示数据呈现正相关的空间聚集。在这种情况下，可以识别出具有高值的聚集/热点区域，即与周围地区相比，这些区域具有更高或更低的观测值。可以通过生成空间统计图表或使用空间分析软件来可视化这些聚集/热点区域。**

**异常值：异常值通常表现为与周围地区明显不同的观测值。在Geary’s C指数的分析中，异常值可能会导致指数的值偏离期望的范围（接近1）。当Geary’s C指数的值远离1时，可以进一步调查与之相关的地理区域，以确定是否存在异常值。这些异常值可能代表了特殊的地理现象或数据采集错误等问题。**

1. **When working with lattice data, there are typically two forms of filters used for the weight matrix. Please points the following filters, and give their spatial weights matrices, respectively. Also, please write a line code for create the weighting matrix according to any one of the weight filters.**

**在处理格网数据时，通常使用两种形式的过滤器用于权重矩阵。请指出以下过滤器，并分别给出它们的空间权重矩阵。此外，请编写一行代码，根据任何一种权重过滤器创建权重矩阵。**



答：

①左边的权重过滤器为Rook（共边）过滤器，Rook过滤器只考虑一个像素的上下左右四个相邻像素，不包括对角线方向。对应的空间权重矩阵是一个对称的二进制矩阵，对角线为0，非对角线元素为1。

其空间权重矩阵为：



这个矩阵表示每个格点与其上下左右相邻的格点之间的连接关系，其中 1 表示相邻，0 表示不相邻。

②右边的权重过滤器为Queen's （共顶点和边）过滤器，Queen's 过滤器考虑一个像素的所有相邻像素（包括对角线方向），形成一个8邻域。对应的空间权重矩阵是一个对称的二进制矩阵，对角线为0，非对角线元素为1。

其空间权重矩阵为：



③根据Rook权重过滤器创建权重矩阵代码：

rook\_filter <- matrix(c(0, 1, 0,

1, 1, 1,

0, 1, 0), nrow = 3, byrow = TRUE)

根据Queen权重过滤器创建权重矩阵代码：

queen\_filter <- matrix(c(1, 1, 1,

1, 1, 1,

1, 1, 1), nrow = 3, byrow = TRUE)