xg1990

地理圈的碼農

克里金(Kriging)插值的原理與公式推導

學過空間插值的人都知道克里金插值，但是它的變種繁多、公式複雜，還有個半方差函數讓人不知所云

本文講簡單介紹基本克里金插值的原理，及其推理過程，全文分為九個部分：

0.引言－從反距離插值說起
1.克里金插值的定義
2.假設條件
3.無偏約束條件
4.優化目標／代價函數
5.代價函數的最優解
6.半方差函數
7.普通克里金與簡單克里金
8.小結

0.引言——從反距離插值(IDW)說起

空間插值問題，就是在已知空間上若干離散點 (x_i, y_i) 的某一屬性(如氣溫，海拔)的觀測值 $z_i = z(x_i, y_i)$的條件下，估計空間上任意一點(x, y)的屬性值的問題。

直觀來講，根據地理學第一定律，

> *All attribute values on a geographic surface are related to each other, but closer values are more strongly related than are more distant ones.*

大意就是，地理屬性有空間相關性，相近的事物會更相似。由此人們發明了反距離插值，對於空間上任意一點(x, y)的屬性$z = z(x, y)$，定義反距離插值公式估計量

$$\hat{z} = \sum_{i=1}^{n} \frac{1}{d^\alpha} z_i$$

其中α通常取1或者2。

即，用空間上所有已知點的數據加權求和來估計未知點的值，權重取決於距離的倒數（或者倒數的平方）。那麼，距離近的點，權重就大；距離遠的點，權重就小。

反距離插值可以有效的基於地理學第一定律估計屬性值空間分佈，但仍然存在很多問題：

- α的值不確定
- 用倒數函數來描述空間關聯程度不夠準確

因此更加準確的克里金插值方法被提出來了

1.克里金插值的定義

相比反距離插值，克里金插值公式更加抽象

$$\hat{z_o} = \sum_{i=1}^{n} \lambda_i z_i$$

其中$\hat{z_o}$是點(x_o, y_o)處的估計值，即$z_o = z(x_o, y_o)$。

這裡的λ_i是權重係數。它同樣是用空間上所有已知點的數據加權求和來估計未知點的值。但權重係數並非距離的倒數，而是能夠滿足點(x_o, y_o)處的估計值$\hat{z_o}$與真實值z_o的差最小的一套最優係數，即

$$\min_{\lambda_i} Var(\hat{z_o} - z_o)$$

同時滿足無偏估計的條件

$$E(\hat{z_o} - z_o) = 0$$

2.假設條件

不同的克里金插值方法的主要差異就是假設條件不同。本文僅介紹普通克里金插值的假設條件與應用。

普通克里金插值的假設條件為，空間屬性z是均一的。對於空間任意一點(x, y)，都有同樣的期望c與方差σ^2。

即對任意點(x, y)都有

$$E[z(x, y)] = E[z] = c$$

$$Var[z(x, y)] = \sigma^2$$

換一種說法：任意一點處的值$z(x,y)$，都由區域平均值c和該點的隨機偏差$R(x,y)$組成，即

$$z(x,y)=E[z(x,y)]+R(x,y)]=c+R(x,y)$$

其中$R(x,y)$表示點(x,y)處的偏差，其方差均為常數

$$Var[R(x,y)]=\sigma^2$$

3.無偏約束條件

先分析無偏估計條件$E(\hat{z_o}-z_o)=0$，將$\hat{z_o}=\sum_{i=1}^n\lambda_i z_i$帶入則有

$$E(\sum_{i=1}^n\lambda_i z_i-z_o)=0$$

又因為對任意的z都有$E[z]=c$，則

$$c\sum_{i=1}^n\lambda_i-c=0$$

即

$$\sum_{i=1}^n\lambda_i=1$$

這是λ_i的約束條件之一。

4.優化目標/代價函數J

再分析估計誤差$Var(\hat{z_o}-z_o)$。為方便公式推理，用符號J表示，即

$$J=Var(\hat{z_o}-z_o)$$

則有

$$\begin{aligned}J&=Var(\textstyle\sum_{i=1}^n\lambda_i z_i-z_o)\\&=Var(\textstyle\sum_{i=1}^n\lambda_i z_i)-2Cov(\textstyle\sum_{i=1}^n\lambda_i z_i,z_o)+Cov(z_o,z_o)\\&=\textstyle\sum_{i=1}^n\sum_{j=0}^n\lambda_i\lambda_j Cov(z_i,z_j)-2\sum_{i=1}^n\lambda_i Cov(z_i,z_o)+Cov(z_o,z_o)\end{aligned}$$

為簡化描述，定義符號 $C_{ij}=Cov(z_i,z_j)=Cov(R_i,R_j)$，這裡$R_i=z_i-c$，即點$(x_i,y_i)$处的属性值相对于区域平均属性值的偏差。

则有

$$J=\sum_{i=1}^{n}\sum_{j}^{n}\lambda_i\lambda_jC_{ij}-2\sum_{i=1}^{n}\lambda_iC_{io}+C_{oo}$$

5.代价函数的最优解

再定义半方差函数 $r_{ij}=\sigma^2-C_{ij}$，带入J中，有

$$\begin{aligned}J&=\sum_{i=1}^{n}\sum_{j=0}^{n}\lambda_i\lambda_j(\sigma^2-r_{ij})-2\sum_{i=1}^{n}\lambda_i(\sigma^2-r_{io})+\sigma^2-r_{oo}\\&=\sum_{i=1}^{n}\sum_{j=0}^{n}\lambda_i\lambda_j(\sigma^2)-\sum_{i=1}^{n}\sum_{j=0}^{n}\lambda_i\lambda_j(r_{ij})-2\sum_{i=1}^{n}\lambda_i(\sigma^2)+2\sum_{i=1}^{n}\lambda_i(r_{io})+\sigma^2-r_{oo}\end{aligned}$$

考虑到$\sum_{i=1}^{n}\lambda_i=1$

$$\begin{aligned}J&=\sigma^2-\sum_{i=1}^{n}\sum_{j}^{n}\lambda_i\lambda_j(r_{ij})-2\sigma^2+2\sum_{i=1}^{n}\lambda_i(r_{io})+\sigma^2-r_{oo}\\&=2\sum_{i=1}^{n}\lambda_i(r_{io})-\sum_{i=1}^{n}\sum_{j=0}^{n}\lambda_i\lambda_j(r_{ij})-r_{oo}\end{aligned}$$

我们的目标是寻找使J最小的一组 λ_i，且J是λ_i的函数，因此直接将J对λ_i求偏导数令其为0即可。即

$$\frac{\partial J}{\partial\lambda_i}=0;i=1,2,\cdots,n$$

但是要注意的是，我们要保证求解出来的最优 λ_i 满足公式$\sum_{i=1}^{n}\lambda_i=1$，这是一个带约束条件的最优化问题。使用拉格朗日乘数法求解，求解方法为构造一个新的目标函数

$$J+2\phi(\sum_{i=1}^{n}\lambda_i-1)$$

其中ϕ是拉格朗日乘数。求解使这个代价函数最小的参数集$\phi,\lambda_1,\lambda_2,\cdots,\lambda_n$，则能满足其在$\sum_{i=1}^{n}\lambda_i=1$约束下最小化$J$。即

$$\begin{cases}\frac{\partial(J+2\phi(\sum_{i=1}^{n}\lambda_i-1))}{\partial\lambda_k}&=0;k=1,2,\cdots,n\\\frac{\partial(J+2\phi(\sum_{i=1}^{n}\lambda_i-1))}{\partial\phi}&=0\end{cases}$$

$$\begin{cases}\frac{\partial(2\sum_{i=1}^{n}\lambda_i(r_{io})-\sum_{i=1}^{n}\sum_{j}^{n}\lambda_i\lambda_j(r_{ij})-r_{oo}+2\phi(\sum_{i=1}^{n}\lambda_i-1))}{\partial\lambda_k}=0;k=1,2,\cdots,n\\\frac{\partial(2\sum_{i=1}^{n}\lambda_i(r_{io})-\sum_{i=1}^{n}\sum_{j}^{n}\lambda_i\lambda_j(r_{ij})-r_{oo}+2\phi(\sum_{i=1}^{n}\lambda_i-1))}{\partial\phi}=0\end{cases}$$

$$\begin{cases}2r_{ko}-\sum_{j=1}^{n}(r_{kj}+r_{jk})\lambda_j+2\phi&=0;k=1,2,\cdots,n\\\sum_{i=1}^{n}\lambda_i&=1\end{cases}$$

由于$C_{ij}=Cov(z_i,z_j)=C_{ji}$，因此同样地$r_{ij}=r_{ji}$，那么有

$$\begin{cases} r_{ko} - \sum_{j=1}^{n} r_{kj}\lambda_j + \phi & = 0; k = 1, 2, \cdots, n \\ \sum_{i=1}^{n} \lambda_i & = 1 \end{cases}$$

式子中半方差函数r_{ij}十分重要，最后会详细解释其计算与定义

在以上计算中我们得到了对于求解权重系数λ_j的方程组。写成线性方程组的形式就是：

$$\begin{cases} r_{11}\lambda_1 + r_{12}\lambda_2 + \cdots + r_{1n}\lambda_n - \phi & = r_{1o} \\ r_{21}\lambda_1 + r_{22}\lambda_2 + \cdots + r_{2n}\lambda_n - \phi & = r_{2o} \\ & \cdots \\ r_{n1}\lambda_1 + r_{n2}\lambda_2 + \cdots + r_{nn}\lambda_n - \phi & = r_{no} \\ \lambda_1 + \lambda_2 + \cdots + \lambda_n & = 1 \end{cases}$$

写成矩阵形式即为

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} & 1 \\ r_{21} & r_{22} & \cdots & r_{2n} & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_n \\ -\phi \end{bmatrix} = \begin{bmatrix} r_{1o} \\ r_{2o} \\ \cdots \\ r_{no} \\ 1 \end{bmatrix}$$

对矩阵求逆即可求解。

唯一未知的就是上文中定义的半方差函数r_{ij}，接下来将详细讨论

6.半方差函数

上文中对半方差函数的定义为

$$r_{ij} = \sigma^2 - C_{ij}$$

其等价形式为

$$r_{ij} = \frac{1}{2}E[(z_i - z_j)^2]$$

这也是半方差函数名称的来由，接下来证明这二者是等价的：

根据上文定义 $R_i = z_i - c$，有$z_i - z_j = R_i - R_j$，则

$$\begin{aligned} r_{ij} & = \tfrac{1}{2}E[(R_i - R_j)^2] \\ & = \tfrac{1}{2}E[R_i^2 - 2R_iR_j + R_j^2] \\ & = \tfrac{1}{2}E[R_i^2] + \tfrac{1}{2}E[R_j^2] - E[R_iR_j] \end{aligned}$$

又因为：

$$E[R_i^2]=E[R_j^2]=E[(z_i-c)^2]=Var(z_i)=\sigma^2$$

$$E[R_iR_j]=E[(z_i-c)(z_j-c)]=Cov(z_i,z_j)=C_{ij}$$

于是有

$$\begin{aligned} r_{ij} &= \tfrac{1}{2}E[(z_i-z_j)^2] \\ &= \tfrac{1}{2}E[R_i^2]+\tfrac{1}{2}E[R_j^2]-E[R_iR_j] \\ &= \tfrac{1}{2}\sigma^2+\tfrac{1}{2}\sigma^2-C_{ij} \\ &= \sigma^2-C_{ij} \end{aligned}$$

$\sigma^2-C_{ij}=\frac{1}{2}E[(z_i-z_j)^2]$得证，现在的问题就是如何计算

$$r_{ij}=\frac{1}{2}E[(z_i-z_j)^2]$$

这时需要用到地理学第一定律，空间上相近的属性相近。$r_{ij}=\frac{1}{2}(z_i-z_j)^2$表达了属性的相似度；空间的相似度就用距离来表达，定义i与j之间的几何距离

$$d_{ij}=d(z_i,z_j)=d((x_i,y_i),(x_j,y_j))=\sqrt{(x_i-x_j)^2+(y_i-y_j)^2}$$

克里金插值假设r_{ij}与d_{ij}存在着函数关系，这种函数关系可以是线性、二次函数、指数、对数关系。为了确认这种关系，我们需要首先对观测数据集

$$\{z(x_1,y_1),z(x_2,y_2),z(x_3,y_3),\cdots,z(x_{n-1},y_{n-1}),z(x_n,y_n)\}$$

计算任意两个点的 距离$d_{ij}=\sqrt{(x_i-x_j)^2+(y_i-y_j)^2}$和 半方差 $\sigma^2-C_{ij}=\frac{1}{2}E[(z_i-z_j)^2]$，这时会得到$n^2$个$(d_{ij},r_{ij})$的数据对。

将所有的d和r绘制成散点图，寻找一个最优的拟合曲线拟合d与r的关系，得到函数关系式

$$r=r(d)$$

那么对于任意两点$(x_i,y_i),(x_j,y_j)$，先计算其距离d_{ij}，然后根据得到的函数关系就可以得到这两点的半方差r_{ij}

7. 简单克里金（simple kriging）与普通克里金（ordinary kriging）的区别

以上介绍的均为普通克里金（ordinary kriging）的公式与推理。

事实上普通克里金插值还有简化版，即简单克里金（simple kriging）插值。二者的差异就在于如何定义插值形式：

上文讲到，普通克里金插值形式为

$$\hat{z_o} = \sum_{i=1}^{n} \lambda_i z_i$$

而简单克里金的形式则为

$$\hat{z_o} - c = \sum_{i=1}^{n} \lambda_i (z_i - c)$$

这里的符号c在上文介绍过了，是属性值的数学期望，即$E[z] = c$。也就是说，在普通克里金插值中，认为未知点的属性值是已知点的属性值的加权求和；而在简单克里金插值中，假设未知点的属性值相对于平均值的偏差是已知点的属性值相对于平均值的偏差的加权求和，用公式表达即为：

$$\hat{R_o} = \sum_{i=1}^{n} \lambda_i R_i$$

这里的R_i在上文定义过了：$R_i = z_i - c$。

但是为什么这样的克里金插值称为“简单克里金”呢？由于有假设$E[z] = c$，也就是说$E(R_i + c) = c$，即$E(R_i) = 0$。那么上面的公式$\hat{R_o} = \sum_{i=1}^{n} \lambda_i R_i$两边的期望一定相同，那么在求解未知参数$\lambda_i$就不需要有无偏约束条件$\sum_{i=1}^{n} \lambda_i = 1$。换句话说，这样的估计公式天生就能满足无偏条件。因此它被称为简单克里金。

从在上文（第4节优化目标/代价函数J）中可以知道，优化目标的推理和求解过程是通过对属性值相对于期望的偏差量R_i进行数学计算而进行的。也就是说这两种克里金插值方法虽然插值形式不一样，求解方法是一样的，重要的区别是简单克里金插值不需要约束条件$\sum_{i=1}^{n} \lambda_i = 1$，求解方程组为：

$$\begin{cases} r_{11}\lambda_1 + r_{12}\lambda_2 + \cdots + r_{1n}\lambda_n + \phi & = r_{1o} \\ r_{21}\lambda_1 + r_{22}\lambda_2 + \cdots + r_{2n}\lambda_n + \phi & = r_{2o} \\ & \cdots \\ r_{n1}\lambda_1 + r_{n2}\lambda_2 + \cdots + r_{nn}\lambda_n + \phi & = r_{no} \end{cases}$$

还有更重要的一点，简单克里金的插值公式为：

$$\hat{z_o} = \sum_{i=1}^{n} \lambda_i (z_i - c) + c$$

换句话说，在计算未知点属性值$\hat{z_o}$前，需要知道该地区的属性值期望c。事实上我们在进行插值前很难知道这个地区的真实属性值期望。有些研究者可能会采用对观测数据简单求平均的方法计算期

望值c，而考虑到空间采样点位置代表性可能有偏差（比如采样点聚集在某一小片地区，没有代表性），简单平均估计的期望也可能是有偏差的。这是简单克里金方法的局限性。

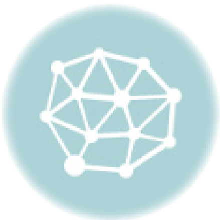
8.小结

总的来说，进行克里金插值分为这几个步骤：

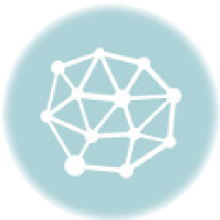
1. 对于观测数据，两两计算距离与半方差
2. 寻找一个拟合曲线拟合距离与半方差的关系，从而能根据任意距离计算出相应的半方差
3. 计算出所有已知点之间的半方差r_{ij}
4. 对于未知点z_o，计算它到所有已知点z_i的半方差r_{io}
5. 求解第四节中的方程组，得到最优系数λ_i
6. 使用最优系数对已知点的属性值进行加权求和，得到未知点z_o的估计值

打赏

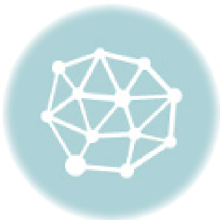
相关文章:



加权线性回归 (Weighted



用**WRF**模型进行气象模拟入门**(1)**



用**WRF**模型进行气象模拟入门**(2)**

2014年10月15日 xg1990 地学、技术 GIS、插值、统计

《克里金(**Kriging**)插值的原理与公式推导》有**78**个想法

yangbaoan

2015年3月14日 18:26