“Slepian-Wolf” 编码及相关性研究

黄冬勃

2015 年 11 月 26 日

目录

Chapter 1

Introduction

SW 编码主要利用信源的空间相关性，要研究一种较为通用是要对信源进行建模，需要知道信源模型所涉及参数相关性，这样，就需要先对这些参数变量的相关性进行建模。

关于时空相关性，要注意并理解清楚时间相关性的切入点，如何和空间相关性联合考虑，或者，它们可以相互独立影响信源参数变化？并且，如何使 SW 编码利用时间相关性？

时间关系也许可以这样使用，假设簇头节点 t_i 时刻观测值变化幅度比上一时刻 t_{i-1} 测量值较大，根据空间相关性，簇内其他节点 t_i 时刻测量值，比其各自 t_{i-1} 两时刻变化幅度也会比较大。找出这个规律，则可以在编码阶段，相比仅利用空间相关性时，保证相同的可靠性的同时，可以获得更大的压缩率。利用时间相关性，前后参数变化大小来选择簇头，变化最大的节点，首先组播作为簇头？

关于变量的具体分布，在信息采集和编码角度来考虑，也许仅需要把握住其变化特点，即不需要知道它到底变成多少，而是知道两个采集点之间的相关性，其中一个点作为边信息采集后，另一点采集的数据已经是变化后的相关的量了，然后利用相关性进行编码即可。这样我将其分析为"可知与不可知问题"，可知代表我们需要知道相关性，不可知代表我们不需要去"推测"具体数值是多少。

还有一个考虑，若瓦斯或者温度涌出点参数值突然增大，扩撒出去各位置浓度是如何变化，需要推出一个公式。

若参数为瓦斯，则其变化规律与涌出点涌出瓦斯浓度，瓦斯蔓延速度有关，涌出点扩散质的数量为 Q，扩散系数为 D，扩散速度为 v，则点 (x, y, z) 点在 t 时刻的瓦斯浓度 C 应为：

$$C = \frac{Q}{2\sqrt{\pi D t}} e^{-\frac{x^2}{4D}}$$

Chapter 2

关于运移形式分析的学习

根据傅里叶 1822 年建立的导热方程，建立定量公式，在 $\triangle t$ 时间内，沿 x 方向通过 x 处截面所迁移的物质的量 $\triangle m$ 与 x 处的浓度成正比

$$\triangle m \propto \frac{\triangle C}{\triangle x} A \triangle t$$

转化为微分形式

$$\frac{dm}{Adt} = -D_m \left(\frac{\partial C}{\partial x} \right)$$

根据扩散通量的概念有：

$$J = -D_m \frac{\partial C}{\partial x} \tag{2.0.1}$$

eq. (2.0.1) 就是菲克第一扩散定律，J 代表扩散通量，单位是 $mol/(cm^2) \cdot s$，D_m 代表扩散质的分子扩散系数，其单位是 cm^2/s 或 m^2/s；$\frac{\partial C}{\partial x}$ 为扩散质的浓度梯度；负号代表扩散质的浓度梯度方向与扩散方向相反。

扩散通量：单位时间垂直通过扩散方向的单位面积的物质的流量。

扩散系数：当浓度为一个单位时，单位时间内，垂直通过扩散方向的物质流量。

菲克第二定律：非稳态状况时，根据 eq. (2.0.1) 不易求出 $C(x,t)$。从物质的平衡关系出发，菲克提出了第二个微分方程：在扩散方向上取体积元 $A\triangle x$，J_x 和 $J_{x+\triangle x}$ 分别表示流入体积元及流出体积元的扩散通量，则在 $\triangle t$ 时间内，体积元中扩散物质的累计量为：

$$\triangle m = (J_x A - J_{x+\triangle x}) \triangle t$$

则有

$$\frac{\triangle m}{\triangle x A \triangle t} = \frac{J_x - J_{x+\triangle x}}{\triangle x}$$

当 $\triangle x, \triangle t > 0$ 时，有：

$$\frac{\partial C}{\partial t} = -\frac{\partial J}{\partial x}$$

将 eq. (2.0.1) 带入上式：

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left(D_m \frac{\partial C}{\partial x} \right) \tag{2.0.2}$$

若扩散系数 D_m 与浓度无关，则式 eq. (2.0.2) 可写为：

$$\frac{\partial C}{\partial t} = D_m \frac{\partial^2 C}{\partial x^2} \tag{2.0.3}$$

一般称 eqs. (2.0.2) and (2.0.3) 为菲克第二定律，其积分解为：

$$C(x,t) = \frac{\mathcal{Q}}{2\sqrt{\pi D_m t}} e^{-\frac{x^2}{4D_m}} \tag{2.0.4}$$

式中 $\mathcal{Q}$ 为 $t = 0$ 时在 $x = 0$ 处的扩散质的数量，这些扩散质沿 x 的方向扩散。该式表示扩散质的浓度 C 沿 x 的分布规律，该浓度按指数规律急剧衰减。

考虑到通风情况，瓦斯的对流运移通量可以表示为：

$$J=\frac{Q_g}{S}=\frac{CQ}{S}=\frac{C\cdot(S\cdot U)}{S}=C\cdot U \tag{2.0.5}$$

其中 U 表示风流的平均速度向量；S 表示风流通过的表面积；$Q=S\cdot U$ 表示通过 S 表面积的风流流量向量；C 表示瓦斯浓度；$Q_g=S\cdot Q$ 表示对流运移通过的瓦斯流量向量。

那么，对于一个测点，也可看做一个微元，其扩散质浓度如何计算？微元内扩散质的变化量可以描述为：

$$\begin{aligned}\delta_Q &= \delta_{Q_{in}}-\delta_{Q_{out}}\\ &= \left(Cudydzdt-D_m\frac{\partial C}{\partial x}dydzdt\right)-\left(\left(Cu+\frac{\partial Cu}{\partial x}dx\right)dydzdt-\left(D_m\frac{\partial C}{\partial x}+\frac{\partial}{\partial x}\left(D_m\frac{\partial C}{\partial x}\right)dx\right)dydzdt\right)\\ &= \frac{\partial}{\partial x}(Cu-D_m\frac{\partial C}{\partial x})dxdydzdt\end{aligned} \tag{2.0.6}$$

由于浓度 C 的变化，在 $\triangle t$ 时间内，微元内扩散质的量的变化 δ_{Q_C} 表示为：

$$\delta_{Q_c}=\frac{\partial C}{\partial t}dtdxdydz \tag{2.0.7}$$

根据质量守恒定律可得：

$$\delta_Q+\delta_{Q_c}=Jdxdydzdt \tag{2.0.8}$$

其中，J 表示单位时间内由于化学、物理等变化导致的单位体积内扩散质的变化量。整理可得瓦斯在巷道内流动的运移扩散方程：

$$\frac{\partial C}{\partial t}+\frac{\partial(Cu)}{\partial x}=D_m\frac{\partial^2 C}{\partial x^2}+J \tag{2.0.9}$$

同时考虑风流运移和分子扩散因素，参考菲克扩散第二定律，可出瓦斯瞬时源在一维湍流扩散作用下巷道瓦斯浓度分布为：

$$C(x,t)=\frac{M}{2\sqrt{\pi D_t t}}exp\left(-\frac{(x-ut)^2}{4D_t t}\right) \tag{2.0.10}$$

其中，$C(x,t)$ 为扩散质浓度，是 t 时间内沿 x 方向扩散质的浓度值，D_t 是瓦斯的湍流扩散系数，u 表示风流速度，M 为 $t=0,x=0$ 时，x 方向扩散质的总量。

以轴对称的圆形巷道中瓦斯运移扩散模型为基础，利用质量传递和动量传递比拟原理。结合巷道湍流风速分布函数，可以得出巷道中瓦斯运移扩散系统 D_t 的表达式：

$$D_t=65.74r\sqrt{\alpha u} \tag{2.0.11}$$

其中，r 为巷道的半径，α 为巷道的摩擦力系数，u 为巷道断面的平均风速。

如果给定的不是原点的浓度，而是连续加入的扩散质的量，而且是变化的，则可以看作无数不同强度的瞬时源产生的扩散瓦斯在时间上的叠加的结果，每个微时间段加入的扩散质所产生的浓度，即

$$\delta_C=\frac{f(\tau)d\tau}{2\sqrt{\pi D_t(t-\tau)}}exp\left\{\frac{|x_1-u_1(t-\tau)|^2}{4D_t(t-\tau)}\right\} \tag{2.0.12}$$

按瞬时源的关系考虑，然后进行从 0 到 t 得时间积分，得：

$$C(x,t)=\int_0^t\frac{f(\tau)}{2\sqrt{\pi D_t(t-\tau)}}exp\left\{-\frac{|x-u_1(t-\tau)|^2}{4D_t(t-\tau)}\right\}dt \tag{2.0.13}$$

Note 1. 已知了扩散质函数，那么考虑参数在各个采集点得分布特征。根据 *eq.* (2.0.4) 和 *eq.* (2.0.10) 可知扩散过程与时间空间均有关系，接下来是否需要考虑，若使用 *SW* 编码，如何综合利用该函数。还需要确定的一个是，公式中的 t 究竟是什么概念，是在 t 时刻还是 t 时间段以内。

关于相关性，若两个变量是非线性的，则可以通过有限元的方法，分段线性化，则变量之间相关性可使用协方差，相关系数来描述。

2.1 空间统计学一些原理

空间统计学是以空间连续性理论和区域化变量理论为基础，以半变异函数为基本工具的一种数学方法。

2.1.1 区域化变量

区域化变量是指以空间点 X 的直角坐标值 (x_u, x_v, x_w) 为自变量的随机场 $Z(x_u, x_v, x_w) = X(X)$，是空间统计学的基本概念。区域化变量具有随机性 (变异性) 和结构性两个特征。结构性是指某些地址参数在点 X 与 $X+h$ 处的数值 $Z(X)$ 与 $Z(X+h)$ 具有某种程度的自相关性[1]。

Note 2. 个人理解：区域化变量使得协方差的使用没有了障碍，即分段线性化，否则变量之间的非线性关系将制约协方差的使用，这种分段线性化体现在区域化变量的“结构性”中。

2.1.2 二阶平稳假设 (Stationary Assumption)

设空间随机场 $Z(s)$ 的空间分布律与站点位置无关，即满足

$$F(s_1, \ldots, s_K; z_1, \ldots, z_K) = F(s_1+h, \ldots, s_K+h; z_1, \ldots, z_K) \tag{2.1.1}$$

对于单变量而言，满足：

$$F(s; z) = F(s+h; z) \tag{2.1.2}$$

如果空间随机场 $Z(s)$ 在任一向量 h 均满足时称为严格平稳。概率分布函数可以通过研究区域内所有数据的累积直方图推断获得。严格平稳性假设要求 $Z(s)$ 的各阶矩均存在且平稳，过于严格，在实际中很难满足。一般情况下，满足 1，2 阶矩存在且平稳即可，因而提出二阶平稳性假设，也称之为弱平稳。二阶平稳要求区域化变量 $Z(s)$ 同时满足一下两个条件：

1. 在整个待研空间区域内变量 $Z(s)$ 的数学期望存在，且等于常数，即满足：

$$E[Z(s)] = E[Z(s+h)] = m \qquad \forall h, \forall s \tag{2.1.3}$$

2. 在这个待研空间区域内变量 $Z(s)$ 的协方差函数存在且平稳，即只依赖于滞后向量 h，而与位置 s 无关，满足：

$$\begin{aligned} & Cov\{Z(s), Z(s+h)\} \\ = \; & E[Z(s)Z(s+h)] - E[Z(s)]E[Z(s+h)] \\ = \; & E[Z(s)Z(s+h)] - m^2 \qquad \forall h, \forall s \\ = \; & C(h) \end{aligned} \tag{2.1.4}$$

协方差不依赖于空间的绝对位置，而依赖于相对位置，即具有空间的平稳不变性。特殊的：

$$Cov\{Z(s), Z(s+0)\} = Var[Z(s)] = C(0) \tag{2.1.5}$$

即方差存在且为常数。

2.1.3 本征假设 (Intrinsic Assumption)

本征假设是比二阶平稳假设更弱的假设，因为在实际研究中协方差函数也有不存在的可能，但变异函数却存在。考虑到该情况，本征假设在二阶平稳假设的基础上进一步放宽了对区域化随机变量的要求 (线性)。

当区域化变量 $Z(s)$ 的增量 $Z(s) - Z(s+h)$ 满足以下两个条件时，称其为满足本征假设 (也称为内蕴假设)：

[1] 空间相关性

1. 在整个研究区域内满足：
$$E[Z(s)-Z(s+h)]=0 \qquad \forall h,\forall s \tag{2.1.6}$$

2. 增量 $Z(s)-Z(s+h)$ 的方差函数存在且平稳 (即不依赖于位置 s)，即满足：
$$\begin{aligned} & Var[Z(s)-Z(s+h)] \\ = \; & E[Z(s)-Z(s+h)]^2-\{E[Z(s)-Z(s+h)]\}^2 \\ = \; & E[Z(s)-Z(s+h)]^2 \qquad \forall h,\forall s \\ = \; & 2\gamma(h) \end{aligned} \tag{2.1.7}$$

从上述定义中可看出本征假设的几个特点：

1. 当满足二阶平稳假设时，一定满足本征假设；
2. 二阶平稳假设本质上是对空间域内的所有区域化变量做要求，而本征假设是对区域化变量的增量做要求，当满足增量的平均值是平稳的条件时，区域化变量不一定是二阶平稳的；
3. 本征假设可以用于协方差函数不存在但半变异函数存在的情况。

2.1.4 空间协方差函数

随机过程 $Z(t)$ 在任意时刻 t_1,t_2 处的两个随机变量 $Z(t_1),Z(t_2)$ 的二阶中心混合矩为此随机过程的协方差函数：
$$C(t_1,t_2)=Cov\left[Z(t_1),Z(t_2)\right]=E\left[Z(t_1)Z(t_2)\right]-E\left[Z(t_1)\right]E\left[Z(t_2)\right] \tag{2.1.8}$$
在二维空间域中，定义随机场在任意空间中的两个点 s 和 $s+h$ 处的两个随机变量 $Z(s)$ 和 $Z(s+h)$ 的二阶中心混合矩为此随机场的自协方差函数：
$$C(s,h)=Cov\left[Z(s),Z(s+h)\right]=E\left[Z(s)Z(s+h)\right]-E\left[Z(s)\right]E\left[Z(s+h)\right] \tag{2.1.9}$$
当 eq. (2.1.9) 中的空间向量 h 取 0 时，协方差函数则变为：
$$C(s,s)=Cov\left[Z(s),Z(s+0)\right]=E\left[Z(s)\right]^2-\{E\left[Z(s)\right]\}^2=Var\left[Z(x)\right] \tag{2.1.10}$$

从定义看，空间协方差函数依赖于空间点 s 和空间向量 h。为了使建模过程中具有统计意义，需要区域化变量 $Z(s)$ 满足二阶平稳假设，本征假设。当满足二阶平稳假设时，空间均值与位置无关，即满足：
$$E[Z(s+h)]=E[Z(s)] \qquad \forall h,\forall s \tag{2.1.11}$$
由于此时协方差函数仅与空间向量 h 有关，而与位置 s 无关，通常将 $C(s,h)$ 写为 $C(h)$，eq. (2.1.9) 转化为：
$$C(h)=C(s,h)=E[Z(s)]^2-\{E[Z(s)]\}^2 \tag{2.1.12}$$

2.1.5 空间半变异函数

空间半变异函数是空间统计学建模的基本工具。上述空间协方差函数可以度量随机变量 $Z(s)$ 和 $Z(s+h)$ 之间的空间相关性，即结构性变化；半变异函数不仅可以描述区域化变量的空间结构性变化，而且能够描述其随机变化，对区域变量之间的空间变异性进行量化。半变异函数记为 $\gamma(s,h)$，定义：
$$\begin{aligned} \gamma(s,h) &= \frac{1}{2}Var[Z(s)-Z(s+h)] \\ &= \frac{1}{2}E[Z(s)-Z(s+h)]^2-\frac{1}{2}\{E[Z(s)]-E[Z(s+h)]\}=^2 \end{aligned} \tag{2.1.13}$$

当满足二阶平稳假设时，根据 eq. (2.1.11)：

$$\gamma(h)=\gamma(s,h)=\frac{1}{2}Var[Z(s)-Z(s+h)] \tag{2.1.14}$$

半变异函数和协方差函数的关系式为：

$$\gamma(h)=sill-C(h) \tag{2.1.15}$$

2.1.6 实验半变异函数及模型

根据 eq. (2.1.14)，通过 n 对站点 $Z(s_i)$ 和 $Z(s_i+h)(i=1,2,\ldots,n)$ 的数值，通过求平均值的方法来获得 $\gamma(h)$ 的值。首先将 n 对站点 s_i 和 s_j 的距离为 h 的所有观测值 $Z(s_i)$ 和 $Z(s_i+h)(i=1,2,\ldots,N^h)$ 看成是 $Z(s)$ 和 $Z(s+h)$ 的 N^h 对实现，其中 N^h 表示距离为 h 的站点对的数量。则计算实验变异函数：

$$\gamma^*(h)=\frac{1}{2N^h}\sum_{i=1}^{N^h}\left[Z(s_i)-Z(s_i+h)\right]^2 \tag{2.1.16}$$

2.2 克里金方法 (Ordinrary Kriging)

它是一种求最优、线性、无偏值内插估计量的方法，通过对每个站点的采样值分别赋予一个权重系数，用加权平均法对待估站点进行估计的方法。几种克里金方法及应用：

1. 在满足二阶平稳或者本征假设条件时可以采用普通克里金方法；
2. 在非平稳或者有趋势存在的情况下采用泛克里金方法；
3. 若进行非线性估计，可以采用析取克里金方法；
4. 当区域化变量服从对数分布式，可用对数克里金方法；
5. 对于系数不规则数据，可采用随机克里金方法。

普通克里金方法基本原理：设 $Z(s)$ 是一个二阶平稳的随机函数，某待估几点 s_0 的变程内的 n 个站点的采样值为 $Z(s_1),Z(s_2),\ldots,Z(s_n)$，则 s_0 点得估计量为：

$$Z^*(s_0)=\sum_{i=1}^{n}\lambda_i Z(s_i) \tag{2.2.1}$$

其中，λ_i 为 $Z(s_i)$ 的权重系数，表示各站点对待估站点的影响大小。普通克里金方法的主要问题是通过确定各个权重系数 λ_i，使估计值 $Z^*(s_0)$ 成为真实值 $Z(s_0)$ 的无偏最优估计。

1. 无偏估计条件

   要使 $Z^*(s_0)$ 成为 $Z(s_0)$ 的无偏估计量，需要满足：

   $$E\left[Z^*(s_0)-Z(s_0)\right]=0 \tag{2.2.2}$$

   由于二阶平稳条件下 $E[Z^*(s_0)]=E[Z(s_0)]=m$，进而

   $$E[Z^*(s_0)]=E\left\{\sum_{i=1}^{n}\lambda_i Z(s_i)\right\}=\sum_{i=1}^{n}\lambda_i E[Z(s_i)]=m\sum_{i=1}^{n}\lambda_i \tag{2.2.3}$$

   若使 $E\left[Z^*(s_0)\right]=E[Z(s_0)]$，需要满足 $m\sum_{i=1}^{n}\lambda_i=m$，即：

   $$\sum_{i=1}^{n}\lambda_i=1 \tag{2.2.4}$$

2. 最优估计条件

最优估计条件是指估计值和样本实际值之间的偏差达到最小，一般采用方差来衡量。估计方差的计算公式为：

$$
\begin{aligned}
E\left\{\left[Z^*(s_0)-Z(s_0)\right]^2\right\} &= \sum_{i=0}^{n}\sum_{j=0}^{n}\lambda_i\lambda_j C(s_i,s_j) \\
&= C(s_0,s_0)-2\sum_{i=1}^{n}\lambda_i C(s_0,s_i)+\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j C(s_i,s_j) \qquad (2.2.5)
\end{aligned}
$$

以上问题转换为带等式条件的寻优问题，在 eq. (2.2.4) 的约束条件下，为了使估计的方差为最小值，通过变换方法，采用**拉格朗日 (lagrange) 乘数法**构建如下函数：

$$
F=E\left\{\left[Z^*(s_0)-Z(s_0)\right]^2\right\}-2\mu\left(\sum_{i=1}^{n}\lambda_i-1\right) \tag{2.2.6}
$$

如果有 $\lambda_i(i=1,2,\ldots,n)$ 的一组数值满足这种要求，那么这组数值就能使下式成立：

$$
\begin{cases}
\dfrac{1}{2}\dfrac{\partial F}{\partial \lambda_i}=\sum\limits_{i=1}^{n}\lambda_i C(s_i,s_j)-\mu-C(s_0,s_j)=0 \\
\dfrac{1}{2}\dfrac{\partial F}{\partial \mu}=\left(\sum\limits_{i=1}^{n}\lambda_i-1\right)=0
\end{cases} \tag{2.2.7}
$$

整理得：

$$
\begin{cases}
\sum\limits_{i=1}^{n}\lambda_i C(s_i,s_j)-\mu=C(s_0,s_j) \qquad j=1,2,\ldots,n \\
\sum\limits_{i=1}^{n}\lambda_j=1
\end{cases} \tag{2.2.8}
$$

eq. (2.2.8) 为 $n+1$ 个方程组成的方程组，其中未知变量为 n 个，此方程组称为普通克里金方程组，其转化为矩阵形式：

$$
[K]\,[\lambda]=[M] \tag{2.2.9}
$$

根据 eq. (2.2.9) 可以求得权重系数 λ，进行普通克里金估计时所得的估计方差最小，计算公式：

$$
\sigma_E^2=C(s_0,s_0)-\sum_{i=1}^{n}\lambda_i C(s_0,s_i)+\mu \tag{2.2.10}
$$

若 $Z(s)$ 只满足本征假设，不满足二阶平稳假设时，可利用半变异函数代替协方差函数。根据协方差函数与半变异函数的关系：$C(h)=C(0)-\gamma(h)$，进而利用半变异函数表示的普通克里金方程组：

$$
\begin{cases}
\sum\limits_{i=1}^{n}\lambda_i\gamma(s_i,s_j)+\mu=\gamma(s_0,s_j) \qquad j=1,2,\ldots,n \\
\sum\limits_{i=1}^{n}\lambda_i=1
\end{cases} \tag{2.2.11}
$$

利用半变异函数标识的普通克里金方差为:

$$
\sigma_E^2=\sum_{i=1}^{n}\lambda_i\gamma(s_0,s_i)-\gamma(s_0,s_0)+\mu \tag{2.2.12}
$$

2.3 时空半变异函数和协方差函数

上述空间统计学方法忽略了时间维度中的关联信息。同空间克里金建模方法一样，需要首先确定半变异函数和协方差的理论模型，称为时空半变异函数和时空协方差函数。

空间半变异函数 $\gamma_s(h)$ 是以空间滞后变量 h 为自变量的函数，与空间位置 s 无关，用来表征同一时间点不同空间点的空间变异结构。空间半变异函数与空间协方差函数 $C_s(h)$ 的关系为：

$$
\gamma_s(h)=C_s(0)-C_s(h) \tag{2.3.1}
$$

时间半变异函数 $\gamma_t(d)$ 是以时间之后 d 为自变量的函数，与时间 t 无关，用来表征同一空间点在不同时间点的时间变异结构。时间半变异函数与时间协方差函数 $C_t(d)$ 的关系为：

$$\gamma_t(d) = C_t(0) - C_t(d) \tag{2.3.2}$$

时空半变异函数 $\gamma_{st}(h,d)$ 是以空间滞后变量 h 和时间滞后变量 d 为自变量的函数，用来表征时空点的变异结构。时空半变异函数与时空协方差函数的关系为：

$$\gamma_{st}(h,d) = C_{st}(0,0) - C_{st}(h,d) \tag{2.3.3}$$

时空协方差函数 $C_{st}(h,d)$ 与空间协方差函数 $C_s(h)$ 和时间协方差函数 $C_t(d)$ 有关，采用积和模型表示两者的关系，具体表示为：

$$C_{st}(h,d) = k_1C_s(h)C_t(d) + k_2C_s(h) + k_3C_t(d) \tag{2.3.4}$$

其中引入系数 k_1, k_2, k_3 是为了确保时空协方差函数 $C_{st}(h,d)$ 的正定性。根据 eq. (2.3.1) 和 eq. (2.3.2) 可以得到：

$$C_{st}(hj,d) = k_1\left[C_s(0) - \gamma_s(h)\right]\left[C_t(0) - \gamma_t(d)\right] + k_2\left[C_s(0) - \gamma_s(h)\right] + k_3\left[C_t(0) - \gamma_t(d)\right] \tag{2.3.5}$$

当 $h = 0$,$d = 0$ 时，根据 eq. (2.3.4)：

$$C_{st}(0,0) = k_1C_s(0)C_t(0) + k_2C_s(0) + k_3C_t(0) \tag{2.3.6}$$

将 eqs. (2.3.5) and (2.3.6) 带入 eq. (2.3.3) 得：

$$\gamma_{st}(h,d) = \left[k_2 + k_1C_t(0)\right]\gamma_s(h) + \left[k_3 + k_1C_s(0)\right]\gamma_t(d) - k_1\gamma_s(h)\gamma_t(d) \tag{2.3.7}$$

若 $\gamma_{st}(0,0) = \gamma_s(0) = \gamma_t(0)$，由 eq. (2.3.7) 可得：

$$\gamma_{st}(h,0) = \left[k_2 + k_1C_t(0)\right]\gamma_s(h) \tag{2.3.8}$$

$$\gamma_{st}(0,d) = \left[k_3 + k_1C_s(0)\right]\gamma_t(d) \tag{2.3.9}$$

假设：$k_2 + k_1C_t(0) = 1$；$k_3 + k_1C_s(0) = 1$，联合 eq. (2.3.6) 可得：

$$\begin{aligned} k_1 &= \left[C_s(0) + C_t(0) - C_{st}(0,0)\right]/C_s(0)C_t(0) \\ k_2 &= \left[C_{st}(0,0) - C_t(0)\right]/C_s(0) \\ k_3 &= \left[C_{st}(0,0) - C_s(0)\right]/C_t(0) \end{aligned} \tag{2.3.10}$$

2.4 时空普通克里金插值方法 (Spatio-temporal Ordinary Kriging Interpolation Method)

时空普通克里金差值的预测公式：

$$Z^*(s_0,t_0) = \sum_{i=1}^{n}\sum_{j=1}^{m}\lambda_{(i,j)}Z(s_i,t_j) \tag{2.4.1}$$

其中，$Z^*(s_0,t_0)$ 是在时空点 (s_0,t_0) 插值的估计值，$Z(s_i,t_j)$ 实在时空点 (s_i,t_j) 的采样值，$i = 1,2,\ldots,n$ 为空间站点，$j = 1,2,\ldots,m$ 为时间采样点，预测过程中的总的样本个数为 $n \times m$，$\lambda(i,j)$ 是各时空点的权重系数。

通过与普通克里金相似的推导方法获得最终计算式：

$$[K][\lambda] = [M] \tag{2.4.2}$$

其中

$$[K]=\begin{bmatrix} \gamma_{11,11} & \cdots & \gamma_{1n,11} & \cdots & \gamma_{11,1m} & \cdots & \gamma_{1n,1n} & 1 \\ \gamma_{21,11} & \cdots & \gamma_{2n,11} & \cdots & \gamma_{21,1m} & \cdots & \gamma_{2n,1m} & 1 \\ \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots \\ \gamma_{n1,m1} & \cdots & \gamma_{n1,m1} & \cdots & \gamma_{n1,1m} & \cdots & \gamma_{nn,1m} & 1 \\ \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots \\ \gamma_{nn,m1} & \cdots & \gamma_{nn,m1} & \cdots & \gamma_{nn,nm} & \cdots & \gamma_{nn,mm} & 1 \\ 1 & \cdots & 1 & \cdots & 1 & \cdots & 1 & 0 \end{bmatrix}$$

$$[\lambda]=\begin{bmatrix} \lambda_{(1,1)} & \cdots & \lambda_{(n,1)} & \cdots & \lambda_{(1,m)} & \cdots & \lambda_{(n,m)} & \mu \end{bmatrix}^T$$

$$[M]=\begin{bmatrix} \gamma_{(01,01)} & \cdots & \gamma_{(0n,01)} & \cdots & \gamma_{(01,0m)} & \cdots & \gamma_{(0n,0m)} & 1 \end{bmatrix}^T$$

其中, $\gamma_{ab,xy}=\gamma(a-b,x-y)$ 是时空点对 (s_a-s_b,t_x-t_y) 的变异函数, μ 是拉格朗日乘数, 将获取的值代入 eq. (2.4.1) 获得预测时空值 $Z^*(s_0,t_0)$。

Chapter 3

temperature

[1]

参考文献

[1] A. Ali, A. Khelil, P. Szczytowski, N. Suri, and Acm, *An Adaptive and Composite Spatio-Temporal Data Compression Approach for Wireless Sensor Networks*, ser. Mswim 11: Proceedings of the 14th Acm International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems. New York: Assoc Computing Machinery, 2011. [Online]. Available: <GotoISI>://WOS:000304070700011