

第四节 数据的检查与规整



知识点



地理空间数据检查的方法

地理空间数据的编辑

地理空间数据的规整

地理空间数据的检查与规整

数据检查概念及内容

地理空间数据检查

数据检查就是以数据的规范性、完整性、正确性为检查原则，对数据的定义、组织、精度以及数据间的关系进行全面的检查，以确定数据是否符合质量要求。



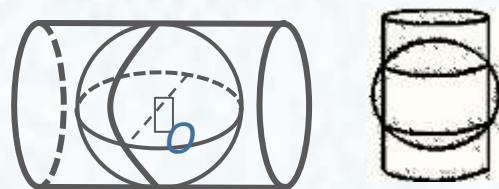
地理空间数据的检查与规整

数据检查——数学基础检查

地理空间数据检查

1. 投影基准

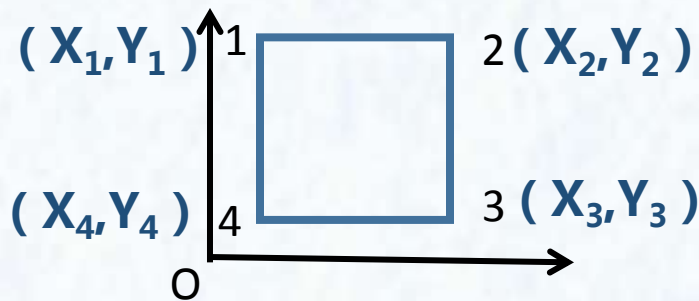
- 在制定检验方案时设定投影基准标准值;
- 通过程序从数据中读出实际投影
- 将两者进行比较,实现自动检查。



投影基准检查

2. 数据范围

- 图廓点顺序是否正确
- 图廓点坐标是否正确
- 数据内容是否在图廓范围内



图廓点顺序图

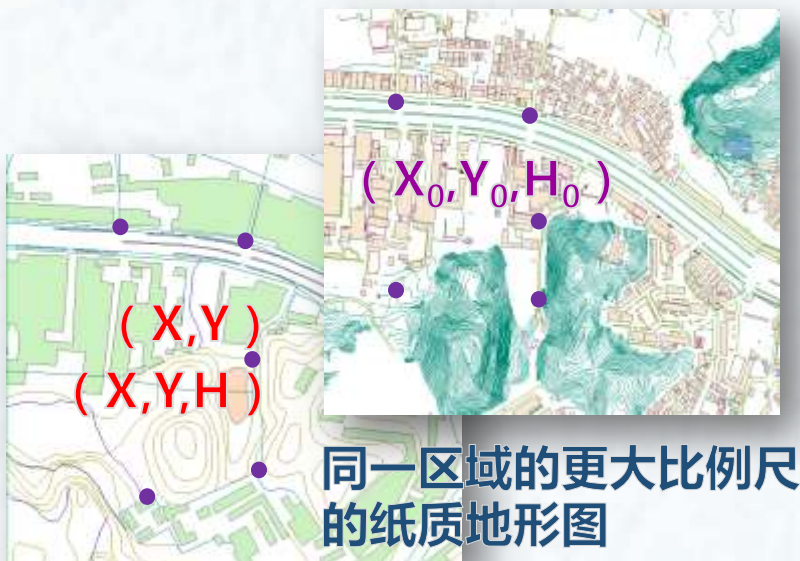
地理空间数据的检查与规整

数据检查——位置检查

位置精度

位置精度包括平面位置精度、高程精度等。

地理空间数据检查



检测步骤：



$$RMSE = \sqrt{(\sum(X - X_0)^2)/N + (\sum(Y - Y_0)^2)/N}$$

$$RMSE_h = \sqrt{(\sum(H - H_0)^2)/N}$$

N为检测点个数

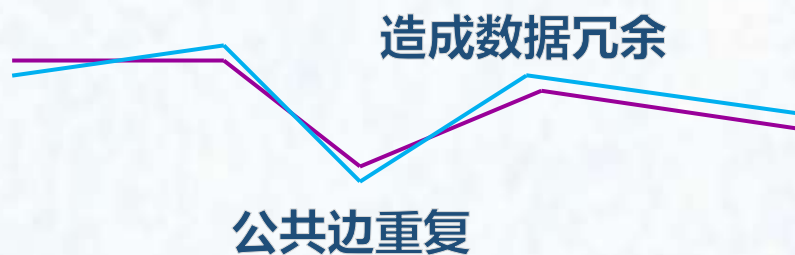
地理空间数据的检查与规整

数据检查——逻辑一致性检查

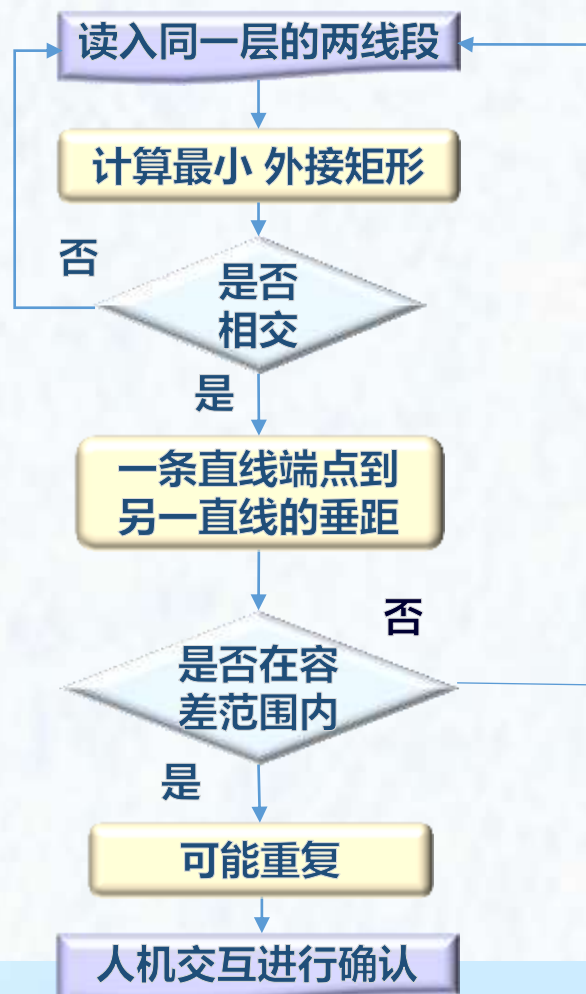
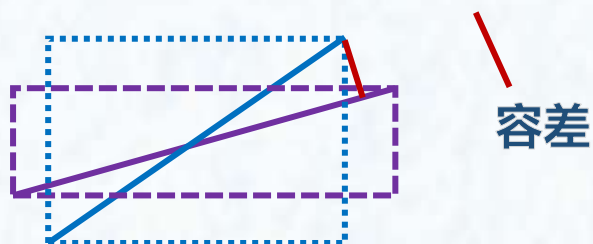
地理空间数据检查

1、重叠性检查

如果同一层内,同类地物的边界被重复输入两次或多次,造成地理空间数据冗余并影响分析的结果。



重叠性检查方法：



地理空间数据的检查与规整

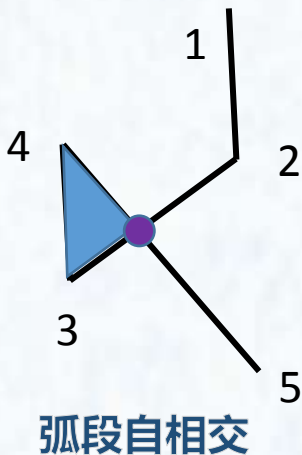
数据检查——逻辑一致性检查

2、弧段相交检查

1) 线段自身相交

同一条折线或曲线自身存在交点。

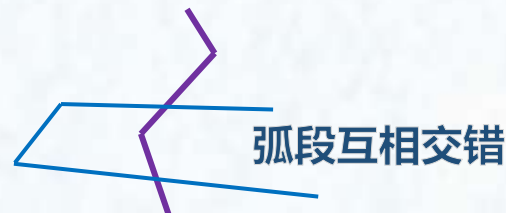
在构建拓扑关系时,会出现无意义的小多边形。



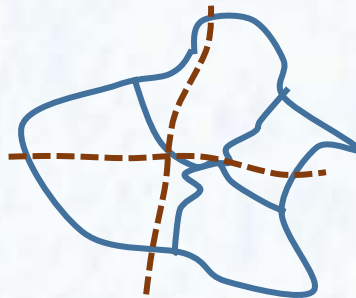
检查方法：对折线相交处求交点，判断是否为端点

2) 弧段互相交错：

两条不应该相交的线存在交点。

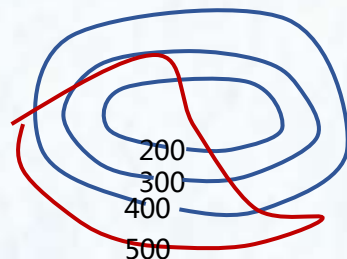


弧段互相交错有合理的情况，但是也有一些不合理的相交。



行政界线

铁路



等高线相交

检查方法：与弧段自相交相同

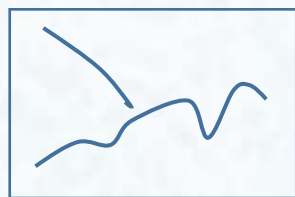
地理空间数据的检查与规整

地理空间数据检查

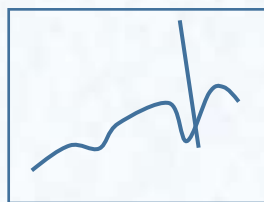
数据检查——逻辑一致性检查

3、悬挂点检查

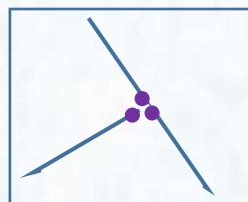
在数据采集过程中, 会产生线段相交不到或相交过头,即存在不合理的悬挂点。



不及

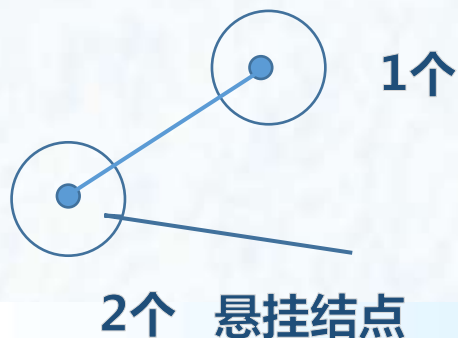


过头



结点不重合

悬挂节点的必要条件:
只与一条线图元相连接



检查方法：



地理空间数据的检查与规整

数据检查——属性数据检查

地理空间数据检查

1、属性项检查

检查方法：

- 属性项数目(5个)
- 属性项定义

2、属性值检查

检查方法：

通过属性值的特性来检查。

- 非法字符检查
- 非空性检查
- 频率法检查
- 固定长度检查
- 属性值范围检查

属性项标准

名称	Not Contain > ? 、
ID	Not null unique
编号	**——**——**
级别	>0
土地利用类型	林地、道路

属性项实例

名称	ID	编号	土地利用类型	级别
>	1234	20-60-12	工业用地	1
黄山	1234	200-60-13	道路	2
?		20-60-14	水池	-3

地理空间数据的检查与规整

地理空间数据检查

数据检查——接边检查

接边检查

1、几何位置接边检查

- 跨图幅的地物在接边线处是否连续
- 根据接边地物的坐标差来衡量接边精度。

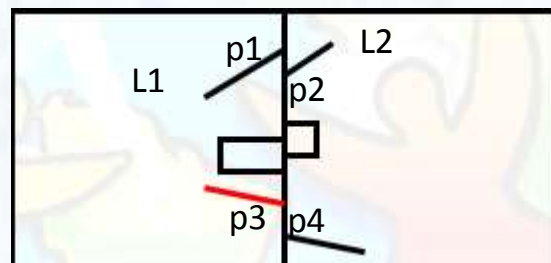
2、属性接边检查

检查在接边线处连续地物的属性是否一致

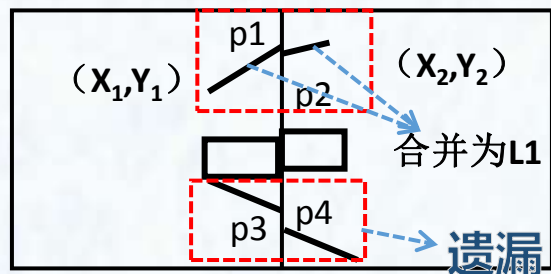
3、接边检查方法:

- ① 寻找接边线
- ② 沿线搜索相邻图幅是否有对应的接边要素
 - 相邻图幅地物遗漏
 - 接边精度是否大于接边限差
- ③ 对应接边要素的属性代码是否一致,

接边前



接边精度： $(X_1 - X_2)^2 + (Y_1 - Y_2)^2$



接边效果

地理空间数据的检查与规整

数据检查——附件质量检查

包括元数据、图历簿、技术设计书及其它文档资料的齐全性、项目填写内容的正确性。

1、手工方法检查

通过逐项对比的方法来进行检测

2、计算机辅助检查

建立“元数据项标准名称模板。

将“标准名称”与“被检元数据项名称”关联起来

通过 取值说明和取值,对元数据文件实现自动检查。

地理空间数据检查



对应标准

控制项

被检元数据项名称	被检元数据项的值	元数据项标准名称	标准名称取值说明	标准名称取值
时间	2013、7	时间信息	采集时间	2012-2014

地理空间数据的检查与规整

数据编辑的目的及内容

地理空间数据编辑

- 目的：
- 减少数字化错误
 - 解决图形接边的问题
 - 减少数据冗余，数据表达不连续等问题

为地理空间数据入库
及应用创造条件



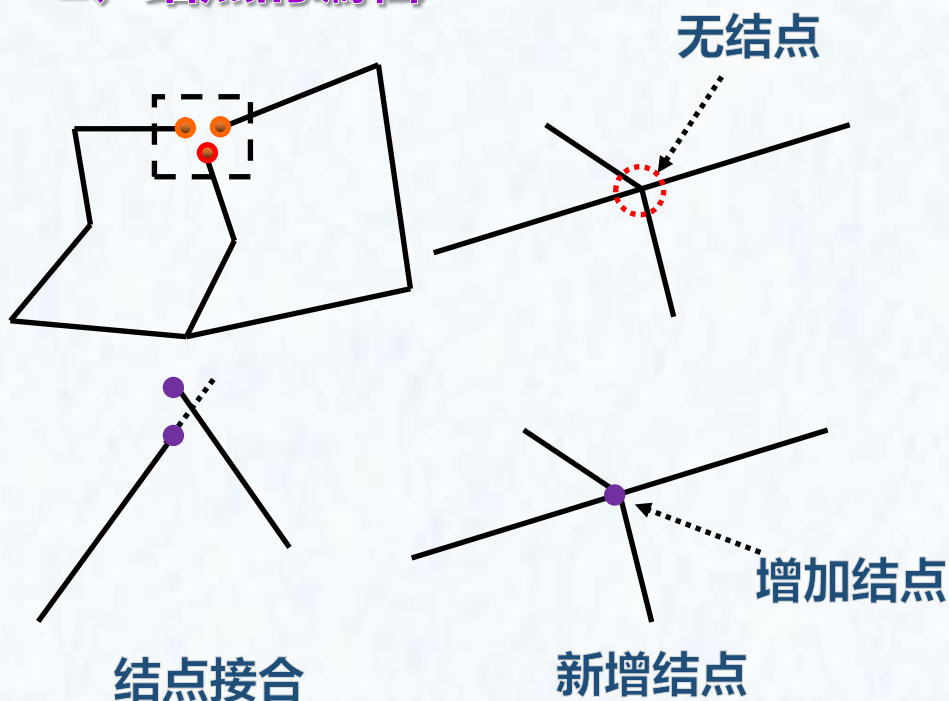
地理空间数据的检查与规整

图形编辑

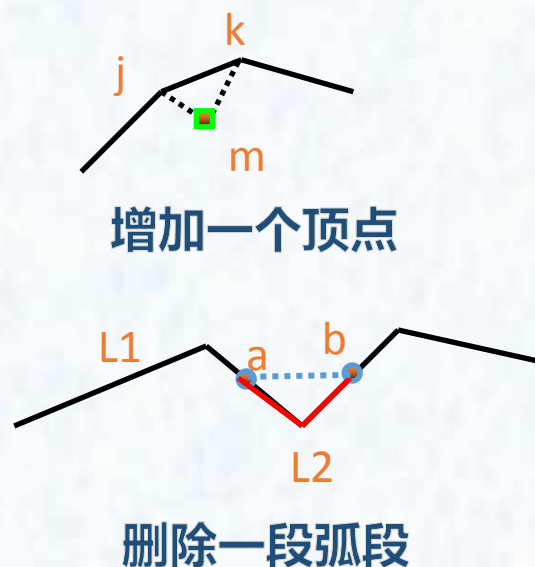
地理空间数据编辑

图形编辑指对矢量数据进行编辑加工，其主要的目的是对几何数据和属性编码进行修改，如点、线、面的增加、删除操作。

1、结点的编辑



2、图形编辑



地理空间数据的检查与整理规整

地理空间数据压缩

地理空间数据编辑

数据检查后，发现存在数据冗余的现象

数据压缩

保证数据的基本特征
简化数据记录
节约存储量

压缩方法

1、矢量数据压缩

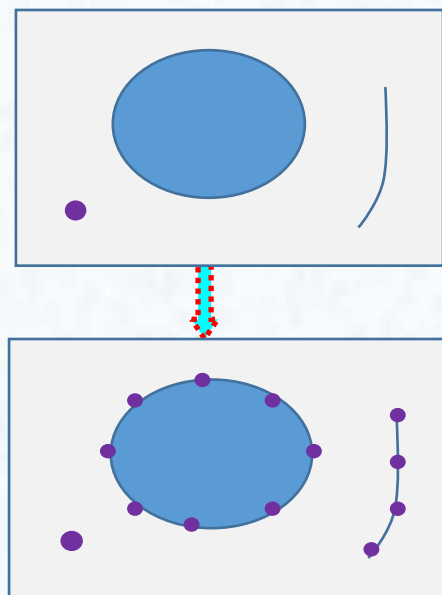
- 道格拉斯普克法
- 垂距法
- 光栏法

2、栅格数据压缩

- 游程编码
- 链状编码
- 四叉树编码

数据压缩的效果

数据压缩



地理空间数据的检查与规整

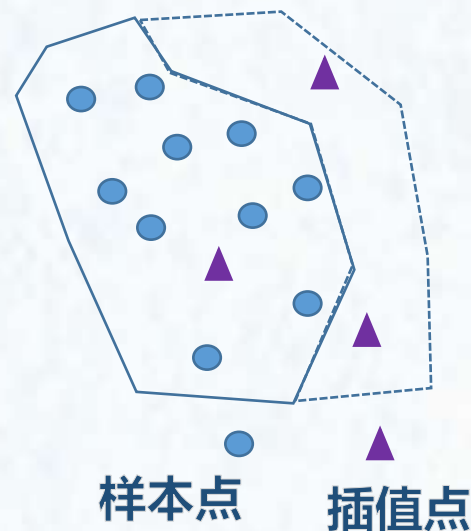
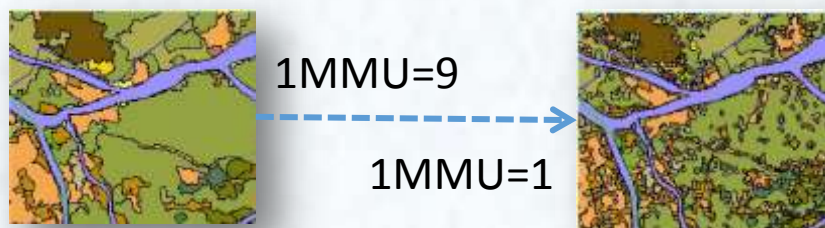
地理空间数据编辑

空间插值

空间插值的原因：

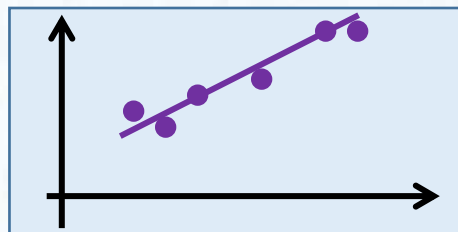
现有数据不能完全覆盖所要求的区域

现有离散曲面的分辨率、像元大小不符合要求

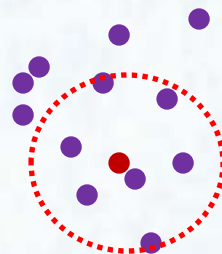


空间插值是在已观测点的区域内或在已观测点的区域外估算未观测点的数据的过程。

空间插值法



趋势面分析法



移动平均法

地理空间数据的检查与规整

数据规整

地理空间数据规整

数据规整的原因：

现有数据往往是不同时期采集的，人工操作多，数据输入操作不严谨。一般存在以下质量问题：

1、地物编码问题

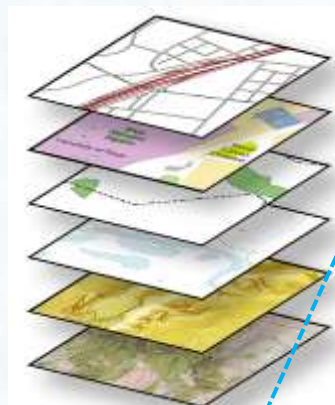
编号	类型	代码
1001	Point	t123
1005	Point	l123
1002	Polygon	
1003	Polyline	L23?
1004	Polygon	g153

编码与符号
线型不一致

缺失
编码

非法
编码

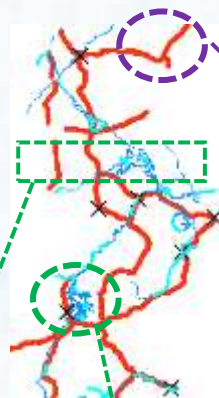
2、地形图分层



街道图
土地利用
线和面状图
水位图
高程图
影像图

垃圾图层

3、要素分层及地物描述问题



地物颜色、符号错误

铁路网

地物要素
分错层

有不明确地物

地理空间数据的检查与规整

数据规整

根据标准规范，对于采集回来的原始数据进行编码赋值，图层划分，格式转换等处理，以达到入库或更新的要求。

数据工具

编码赋值



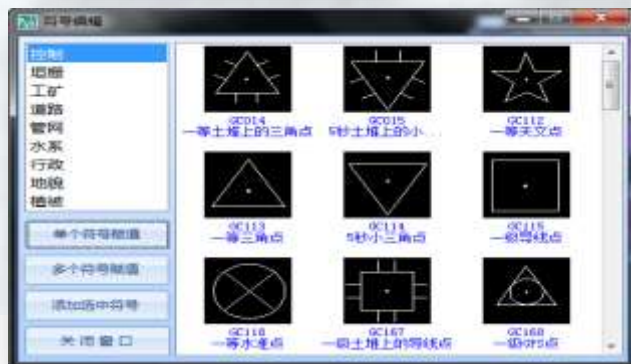
面筛选赋值



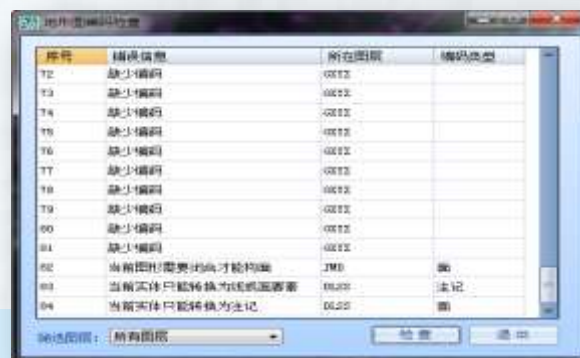
图层筛选



符号编辑



编码检查





谢谢大家！

