

# GIS 中衡量位置数据不确定性的 可视化度量指标族探讨

戴洪磊<sup>1)</sup> 夏宗国<sup>2)</sup> 黄杏元<sup>1)</sup>

<sup>1)</sup> (南京大学城市与资源学系地理信息系统教研室, 南京 210093)

<sup>2)</sup> (麻省大学地理系地理信息系统教研室, 美国波士顿)

**摘要** 点是矢量 GIS 中最基本的元素, 线和面可以由其本身的特征点来加以抽象和表达. 针对当前国际 GIS 数据质量标准中, 只定义了衡量点位数值精度的情况, 应 GIS 可视化的要求, 从图形角度出发, 借鉴测绘学科中衡量平面点元位置不确定性的几何误差椭圆方法, 通过计算点位坐标落入其本身误差椭圆内的概率大小, 来定义矢量 GIS 中衡量点、线和面位置数据不确定性的点位误差椭圆、线位误差带和面位误差环等可视化度量指标族, 以期丰富和完善当前 GIS 数据质量的指标体系, 从而可以更好地控制各种矢量在 GIS 数据库中的精度.

**关键词** 误差椭圆 误差带 误差环 可视化

**中图法分类号:** P208 **文献标识码:** A **文章编号:** 1006-8961(2002)02-0165-05

## Analysing on Visual Indexes of Positional Uncertainty in GIS

DAI Hong-lei<sup>1)</sup>, XIA Zong-guo<sup>2)</sup>, HUANG Xing-yuan<sup>1)</sup>

<sup>1)</sup> (Research Group of Geographic Information Technology, Dept. of Urban Resources, Nanjing University, Nanjing 210093)

<sup>2)</sup> (Research Group of Geographic Information Technology, Dept. of Geography, University of Massachusetts Boston U. S. A)

**Abstract** We know that Line and polygon can be expressed by their feature points because point is the fundamental feature in vector GIS. Because only digital accuracy of point data is defined in the International Geographic Information System (GIS) data quality standard system, so this paper defines the visual indexes of point ellipse, line error band and polygon error donut to assess effect scope of their positional uncertainty according to the probability that the feature points dropping into their error ellipses based on the error ellipse expressing the positional uncertainty in surveying and mapping, we wish to richen and perfect the system of GIS data quality standard system and to control the data quality of many GIS databases to a higher accuracy.

**Keywords** Error ellipse, Error band, Error donut, Visual

## 0 引言

可视化的基本含义是把一些数据转换成图形的形式, 并以视觉的方式去研究这些图形及其变化, 进而获取新知识的一种科技手段<sup>[1]</sup>. 通常可视化技术可以弥补静态的, 以文本为媒介的传输信息方式的不足, 对数据进行实时地跟踪、引导和处理, 从而使

观察者能更容易、更直观、更自然地解释、理解由各种方式获取的大量具有时态变化的地理对象数据. 目前在国内外 GIS 数据质量标准中, 仍只是定义了衡量点位精度的数值度量标准<sup>[2]</sup>, 本文借鉴测绘学科中衡量平面二维点位置数据精度的误差椭圆方法<sup>[3]</sup>, 通过计算随机线元和面元的特征点落入其本身误差椭圆内的概率大小, 来定义 GIS 中用于衡量位置数据不确定性的可视化指标族.

基金项目: 国家自然科学基金(49801016, 40101022)

收稿日期: 2000-10-13; 改回日期: 2001-05-28

# 1 平面随机点的可视化度量指标

## 1.1 平面随机点的误差椭圆

由测绘统计学理论可知,用平行于  $x_1, x_2$  的平面去截二维随机点的概率分布曲面,并将截线投影到平面  $x_1, x_2$  上时,得到一族同心误差椭圆(见图1),其中心为  $(u_1, u_2)$ , 方程为<sup>[3]</sup>

$$\frac{(x_1 - u_1)^2}{\sigma_1^2} - 2\rho \frac{(x_1 - u_1)(x_2 - u_2)}{\sigma_1\sigma_2} + \frac{(x_2 - u_2)^2}{\sigma_2^2} = k_1^2 \quad (1)$$

其中,  $k_1$  为一常数,经过坐标系平移和旋转后,可将上式转化为下列标准形式<sup>[4]</sup>

$$\frac{\bar{x}_1^2}{a_1^2} + \frac{\bar{x}_2^2}{a_2^2} = k_1^2 \quad (2)$$

当  $k_1$  取不同值时,就得到一族同心椭圆,记为  $G_{k_1}$ .

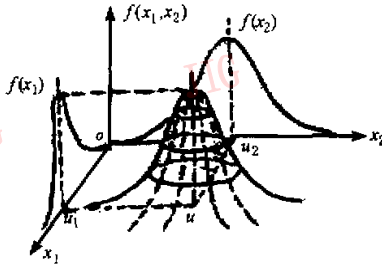


图1 二维正态分布密度图象

经过上述坐标系变换后,二维正态分布密度函数简化为

$$f(\bar{x}_1, \bar{x}_2) = \frac{1}{2\pi a_1 a_2} \exp\left[-\frac{1}{2}\left(\frac{\bar{x}_1^2}{a_1^2} + \frac{\bar{x}_2^2}{a_2^2}\right)\right] \quad (3)$$

根据此式,可以计算随机点  $(x_1, x_2)$  落入误差椭圆  $G_{k_1}$  内的概率,即

$$P\{(x_1, x_2) \in G_{k_1}\} = \iint_{G_{k_1}} f(\bar{x}_1, \bar{x}_2) d\bar{x}_1 d\bar{x}_2 = \iint_{G_{k_1}} \frac{1}{2\pi a_1 a_2} \left(\frac{\bar{x}_1^2}{a_1^2} + \frac{\bar{x}_2^2}{a_2^2}\right) d\bar{x}_1 d\bar{x}_2 \quad (4)$$

令  $y_i = \frac{\bar{x}_i}{\sqrt{2} a_i} (i=1, 2)$ , 代入式(2), 整理得

$$y_1^2 + y_2^2 = \frac{1}{2} k_1^2 \quad (5)$$

上式是以  $\frac{1}{\sqrt{2}} k_1$  为半径的圆  $C_{k_1}$  的方程. 由式(3)和式(5)可知, 随机点  $(x_1, x_2)$  落入椭圆  $G_{k_1}$  内的概率等于落入圆  $C_{k_1}$  内的概率, 称圆  $C_{k_1}$  为等概率圆. 于是式(4)可变为

$$P\{(x_1, x_2) \in G_{k_1}\} = P\{(x_1, x_2) \in C_{k_1}\} = \frac{1}{\pi} \iint_{C_{k_1}} e^{-\frac{1}{2}(y_1^2 + y_2^2)} dy_1 dy_2 = 1 - e^{-\frac{1}{2} k_1^2} \quad (6)$$

对于不同的  $k_1$  值, 利用式(6)可以方便地计算出随机点  $(x_1, x_2)$  落入误差椭圆  $G_{k_1}$  内的概率(表1).

表1 平面点位、线位特征点和面位质点落入其本身误差椭圆内的概率

$k_1$	$P_1$	$P_2$	$P_3$
0	0	0	0
0.5	0.1175	0.2212	0.3127
0.680	0.2064	0.3702	0.50
0.833	0.2931	0.50	0.6468
1.0	0.3935	0.6321	0.7769
1.1774	0.50	0.7500	0.8750
1.24	0.5364	0.7851	0.90
1.4142	0.6321	0.8646	0.9502
1.5	0.6753	0.8946	0.9658
1.518	0.6840	0.90	0.9685
2.0	0.8647	0.9817	0.9975
2.1460	0.90	0.990	0.9990
2.5	0.9560	0.9981	0.99991
3.0	0.9889	0.99988	0.999998
3.5	0.9978	0.999995	0.9999998

## 1.2 平面随机点的椭圆指标

在GIS数据质量标准中,平面坐标点的精度常用圆误差数值指标  $\sigma_c = 0.707 \sqrt{\sigma_{x_1}^2 + \sigma_{x_2}^2} \approx \frac{1}{2}(\sigma_{x_1} + \sigma_{x_2})$  来度量<sup>[5]</sup>, 此外还有圆标准差(CSE)、圆或然差(CPE)、地图圆精度(CMAS)、圆极限差(CNCE)和圆均方差(RMSE)等<sup>[6]</sup>, 它们通常作为二维随机点位置不确定性的度量指标, 是标量值. 误差椭圆作为一种几何图形, 满足GIS可视化的要求, 基于表1中的不同取值, 通过对传统圆误差指标的扩展, 文献[7]提出并定义了以下椭圆指标:

### (1) 标准误差椭圆(SEE)

当  $k_1=1$  时, 称其图象为标准误差椭圆, 简称为误差椭圆, 即众所周知的 Helmert 误差椭圆, 它对应着 39.35% 的椭圆概率.

### (2) 或然误差椭圆(PEE)

当  $k_1=1.1774$  时, 称其几何图形为或然误差椭圆, 对应着 50% 的椭圆概率.

### (3) 地图圆精度椭圆(MCE)

当  $k_1=2.1460$  时, 称其几何图形为地图圆精度

椭圆,对应着 90%的椭圆概率。

(4) 极限误差椭圆(UEE)

当  $k_1=3.5$  时,称其几何图形为极限误差椭圆,对应着 99.78%的椭圆概率,它可用于粗差的初步判别。

(5) 均方误差椭圆(EEE)

当  $k_1=1.4142$  时,称其几何图形为均方误差椭圆,对应着 63.21%的椭圆概率。

上述椭圆族指标的可视化图形如图 2 所示。显然,椭圆指标是对传统圆误差指标的扩展。后者是根据服从圆分布的随机变量  $\sqrt{y_1^2+y_2^2}$  定义的,由  $\sigma_c$  又可进一步定义 CPE,CMAS,CNCE,和 RMSE 指标。为了加以比较,表 2 列出了圆误差指标和椭圆指标。

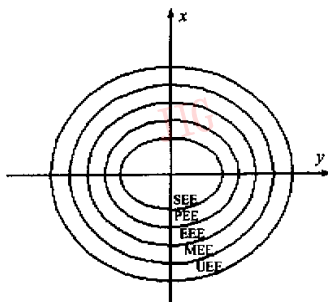


图 2 点位误差椭圆族的可视化

表 2 圆误差指标和椭圆指标

圆误差指标		椭圆指标	椭圆概率 (%)
名称与符号	利用 $\sigma_c$ 定义		
圆标准差 CSE	$CSE=1.0\sigma_c$	标准误差椭圆	39.35
圆或然误差 CPEP	$CPEP=1.1774\sigma_c$	或然误差椭圆	50
地图圆精度 CMAS	$CMAS=2.1460\sigma_c$	地图圆精度椭圆	90
圆极限误差 CNCE	$CNCE=3.5\sigma_c$	极限误差椭圆	99.78
圆均方误差 RMSE	$RMSE=1.4142\sigma_c$	均方误差椭圆	63.21

## 2 随机线元误差带的可视化度量指标

### 2.1 线元上特征点落入其误差椭圆内的概率

由文献[8]的研究可知,线元位置不确定性的影响范围是两端大、中间小的哑铃形条带,即线元上任何一点的精度均高于线元两端点的精度(如图 3),其中精度最高点称为线元上的特征点,其精度解析公式为

$$\sigma_{t_r}^2 = [(1-t_r)^2 + t_r^2](\sigma_{z_0}^2 \cos^2 \varphi + \sigma_{z_1}^2 \sin^2 \varphi + \sigma_{z_2}^2 \sin^2 \varphi) \quad (7)$$

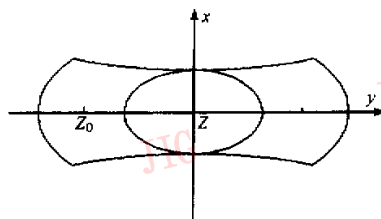


图 3 线元特征点的误差椭圆

其中,  $t_r$  为特征点在线元上的位置,其计算式为:

$$t_r = \frac{\sigma_{z_1}^2 + \sigma_{z_0}^2}{(\sigma_{z_0}^2 + \sigma_{z_1}^2) + (\sigma_{z_1}^2 + \sigma_{z_2}^2)} = \frac{\sigma_{z_0}^2}{\sigma_{z_0}^2 + \sigma_{z_2}^2} \quad (8)$$

为讨论方便计,当假设线元两端点的坐标观测值精度相同(即  $\sigma_{z_0}^2 = \sigma_{z_1}^2$ ),且相互独立时,有  $t_r = 0.5$ ,即特征点位于线元中点,并且其精度  $\sigma_{t_r}^2 = \frac{1}{2}\sigma_{z_0}^2$  为端点精度的一半,将其转化为标准方差

$$\sigma_{t_r}^2 = \frac{1}{2}\sigma_{z_0}^2 = \frac{1}{2}(k_1\sigma_0)^2 = \left(\frac{k_1}{\sqrt{2}}\sigma_0\right)^2 = (k_2\sigma_0)^2 \quad (9)$$

即  $k_1 = \sqrt{2}k_2$ ,将  $k_2$  代入式(6),则精度最高点落入它本身误差椭圆的概率为

$$P\{t_r \in G_{k_1}\} = 1 - e^{-\frac{1}{2}k_1^2} = 1 - e^{-k_2^2} \quad (10)$$

对于不同的  $k_2$  值,利用式(10)可计算出线元特征点  $t_r$  落入误差椭圆  $G_k$  内的概率(见表 1)。

### 2.2 线位误差带指标

仿误差椭圆指标族的定义,根据  $k_2$  的不同取值,可以定义以下线位误差带指标:

(1) 标准差带(SEB)

当  $k_2=1$  时,称线位误差带为标准差带,简称为误差带,它对应着 63.21%的误差带概率。

(2) 或然差带(PEB)

当  $k_2=0.833$  时,称线位误差带为或然差带,它对应着 50.0%的误差带概率。

(3) 地图精度误差带(MEB)

当  $k_2=1.518$  时,称线位误差带为地图  $\epsilon$  精度误差带,对应着 90.0%的误差带概率。

(4) 极限差带(UEB)

当  $k_2=3.0$  时,称线位误差带为极限差带,对应着的 99.988%的误差带概率。

(5) 均方误差椭圆(EEB)

当  $k_2=1.4142$  时,称线位误差带为均方误差椭圆,对应着 86.46%的椭圆概率。

上述线位误差带指标的可视化图形如图4所示。

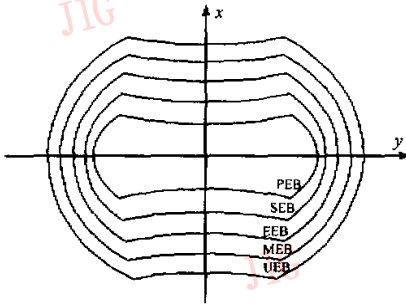


图4 线位误差带族的可视化

### 3 随机面位误差环的可视化度量指标

#### 3.1 面元特征点落入其误差椭圆的概率

对于随机面元,由于它的位置不确定性是由组成它的一系列边界线的位置不确定性来表达的,因此相对单一线元来讲比较复杂,在此限定组成面元的各顶点精度相同且相互独立,故所讨论的随机面元为一匀质随机面元,一般情况下,它可用其特征点质心来代表(如图5所示),其质心公式为<sup>[5]</sup>

$$x_m = \frac{1}{n} \sum_{i=1}^n x_i, y_m = \frac{1}{n} \sum_{i=1}^n y_i \quad (11)$$

根据协方差传播律,可导出质心坐标的方差计算式为

$$\begin{aligned} \sigma_{P_m}^2 &= 2 \sum_{i=1}^n (c_i^2 \sigma_{x_i}^2 + d_i^2 \sigma_{y_i}^2) \\ &= \frac{2}{n^2} \sum_{i=1}^n (\sigma_{x_i}^2 + \sigma_{y_i}^2) = \frac{2}{n^2} \sum_{i=1}^n \sigma_i^2 \end{aligned} \quad (12)$$

其中,  $c_i = d_i = \frac{1}{n}$  为质心坐标  $P_m(x_m, y_m)$  对交点多边形各顶点坐标  $(x_i, y_i)$  的偏导数。通常在GIS中所讨论的匀质多边形边数在3~9之间,考虑边数为6的

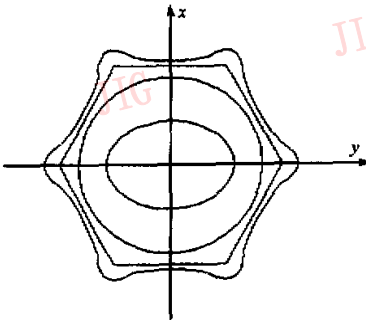


图5 面元特征点的点位误差椭圆

特殊情况,当  $n=6$ ,并将特征点质点的精度转化为标准方差后,式(12)简化为

$$\sigma_{P_m}^2 = \frac{1}{18} \sum_{i=1}^n \sigma_i^2 = \frac{1}{3} \sigma_i^2 = \left( \frac{1}{\sqrt{3}} k_3 \sigma_c \right)^2 = (k_3 \sigma_3)^2 \quad (13)$$

此时,  $k_1 = \sqrt{3} k_3$ ,将  $k_3$  代入式(6),则质点落入它本身误差椭圆的概率计算式为

$$P\{(x_m, y_m) \in G_{k_3}\} = 1 - e^{-\frac{1}{2} k_3^2} = 1 - e^{-\frac{1}{2} k_3^2} \quad (14)$$

对于不同的  $k_3$  值,利用式(14)可以方便地计算出质点落入它的误差椭圆内的概率(见表1)。

#### 3.2 面位误差环度量指标

下面利用质点落入其误差椭圆的概率大小来定义随机面元位置不确定性的可视化度量指标族。

(1) 标准误差环(SED)

当  $k_3=1$  时,称其图象为标准误差环,它对应着 77.69% 的椭圆概率。

(2) 或然误差环(PED)

当  $k_3=0.68$  时,称其几何图形为或然误差环,对应着 50% 的椭圆概率。

(3) 地图圆精度环(MCD)

当  $k_3=1.24$  时,称其几何图形为地图圆精度环,对应着 90% 的椭圆概率。

(4) 极限误差环(UED)

当  $k_3=2.5$  时,称其几何图形为极限误差环,对应着 99.9998% 的椭圆概率。

(5) 均方误差环(EED)

当  $k_3=1.4142$  时,称其几何图形为均方误差环,对应着 95.02% 的椭圆概率。

上述面位误差环指标的可视化图形如图6所示。

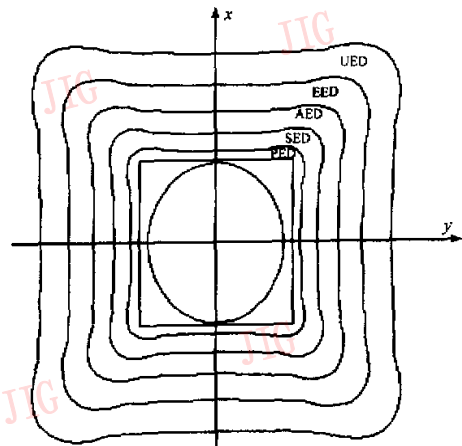


图6 面位误差环族的可视化

### 4 实例

本例选取数据库中一段街区(图 7),采用本文提出的可视化度量指标来衡量它们的精度.对于图中的 3 个区域和一条街道,这里分析街道边 89 和区域 2 365 的精度.因为在图形数字化时,普遍认为取点精度一致,且各点之间相互独立,所以经误差评定,可得数据库中各点的方差值为  $1.29\text{cm}^2$ ,其可视化精度如表 3 所示.

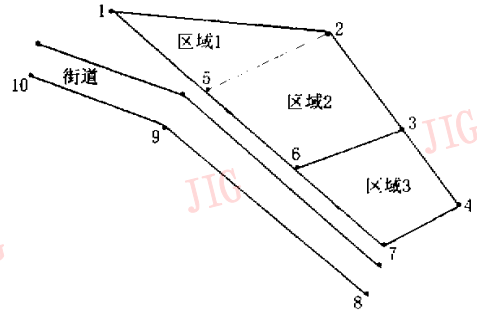


图 7 数据库街区图

表 3 精度评价

	点 5					线 89					面 2365				
	SEE	PEE	MCE	UEE	EEE	SEB	PEB	MEB	UEB	EEB	SED	PED	MCD	UED	EED
可视化精度	1.29	1.52	2.769	4.515	1.824	1.29	1.075	1.958	3.87	1.824	1.29	0.877	1.599	3.225	1.824

### 5 结论

显然,用于衡量 GIS 中线位和面位位置不确定性的可视化指标族,如线位误差带族和面位误差环族是由误差椭圆指标族直接扩展而来的,为了比较,

表 4 列出了相应的 3 种误差指标族.

表 5 中 3 种指标族共同构成了 GIS 中用于衡量点、线和面位置不确定性的可视化指标族.在评价 GIS 位置数据的不确定性时,可通过绘制误差椭圆族和误差带族误差环族来分别度量平面点、线和面的定位不确定性.

表 4 误差椭圆指标、误差带指标和误差环指标族

误差椭圆指标			误差带指标			误差环指标		
名称	k 值	误差椭圆概率(%)	名称	k 值	误差带概率(%)	名称	k 值	误差环概率(%)
标准差椭圆	1.0	39.35	标准差带	1.0	63.21	标准差环	1.0	77.69
或然差椭圆	1.1774	50	或然差带	0.833	50.0	或然差环	0.68	50.0
地图圆精度椭圆	2.1460	90	地图线精度带	1.518	90.0	地图线精度环	1.24	90.0
极限差椭圆	3.0	98.89	极限差带	3.0	99.988	极限差环	2.5	99.9998
均方差椭圆	1.4142	63.21	均方差带	1.4142	86.46	均方差环	1.4142	95.02

### 参考文献

- 郭达志, 盛业华, 余光平等. 地理信息系统基础与应用[M]. 北京: 煤炭工业出版社, 1997. 5.
- 阎正, 蒋景融. 城市地理信息系统标准化指南[M]. 北京: 科学出版社, 1998.
- 於宗伟, 鲁林成. 测量平差基础[M]. 北京: 测绘出版社, 1983.
- 数学手册编写组. 数学手册[M]. 北京: 高等教育出版社, 1984.
- Goodchild M F. Keynote address [A]. In: Edited by Gary Hunter, proceedings of the symposium on spatial database accuracy[C], The University of Melbourne, Australia, 1991: 1~16.
- CCOSAM(Canadian Council on Surveying and Mapping). National Standards for the Exchange of Digital Topographic Data [S], I-Standards for the Quality Evaluation of Digital Topographic Data, 1982.
- 刘文宝. GIS 空间数据的不确定性理论[D]. 武汉: 武汉测绘科技大学, 1995.

- 戴洪磊, 吴守荣, 陈兰森. 矢量 GIS 中线位不确定性的几何象表达[J]. 中国图象图形学报, 1999, 4(12): 1078~1081.



**戴洪磊** 南京大学城市与资源学系博士后, 1998 年获山东科技大学硕士学位, 2000 年获武汉测绘科技大学博士学位. 现主要从事图象处理和 GIS 数据质量分析等方面的研究工作. 已发表论文 20 余篇.

**夏宗国** 美国波士顿麻省大学地理系教授, 博士生导师, GIS 研究中心主任. 主要从事图象、信息与数据处理等方面的研究. 发表论文多篇.

**黄杏元** 教授, 博士生导师, 1961 年毕业于南京大学地理系, 现在南京大学城市与资源学系任教. 主要从事地理信息系统网络和图象处理等方面的工作. 获国家教委科技进步奖多项, 发表论文多篇.