

计算机生产的无级双变量等值区域地图

斯蒂芬·拉夫英及J·克拉克·阿恰

等值区域地图在地理分布的分析方面通常用于地图模式的目视比较。正因为如此,大量的制图研究者集中于面状分布吻合或相关程度的视觉传输的处理上(登脱,1965;……)。通常这些研究涉及到地图读者识别等值区域地图模式之间相似性或非相似性的能力,而这某等值区域地图是根据表示单变量分布的统计分级经符号化而得到的(穆勒,1976)。

单变量场合下的等值区域地图的制图技术已是一种多年来的规范化实践。然而,最近的研究启示人们,无级等值区域地图和双变量等值区域地图都成为有效的传输工具。本文另外提出了一种传达两种地理分布之间面状吻合程度的技术,它采用了在线划绘图机或CRT上产生无级双变量等值区域地图。除了考察这类地图的分析及专题背景以外,我们也要谈到为生产这类地图而研制的计算机程序——BICHOR的应用。

究竟要不要量子化 在地图设计中,人们早已采用数据分级方法来表示经过综合的地理分布形式。在1973年,这个传统方法受到了托布勒有创见地挑战。他提出不用分级间隔产生等值区域图,而用计算机驱动的线划绘图机以实际上为连续的灰色调生产出地图。他坚信认为地图读者有能力区分出地图上大量不同的灰色调,这正如他们有能力目视处理航空像片上的连续色调一样。托布勒很有创见地观察到,无级地图避免了使用分级间隔中固有的量子化误差问题。这一点很快就引起了常规知识捍卫者们的反应。反对无级地图的论据,主要依赖于人类的视觉处理系统在对信息的编码和贮存中的天生的局限性。这样,分级等值区域图比起无级等值区域地图来,在更大程度上考虑了人类处理机能的范围(道不森,1973,……)。

之所以会产生制图上的意见分歧,部分原因是基于对地图功能的认识这一哲学上的差异。托布勒的研究重点是分析制图,或者说将地图作为系统的阐述假设、验证假设,并且最终解释地理分布的一种工具(托布勒,1976)。分析制图主张,人们不光是简单地看看地图,而且要整体地及逐区地研究地图,以求获得地理上的解释。相反,地图传输所关心的是地理图解,其目的是以最小的感受误差快速传输一种分布的醒目的视觉特征。

彼得森在用分级与无级的等值区域图作地图对照的实验研究中,发现“在无级等值区域图上所描述的富余信息并不妨碍在这些地图之间所作的对照比较”(彼得森,

应用研究,然而不应得出这样的推论,即概率法则 $\pi(x)$ 只局限于城市研究的有关问题。概率法则是对形成完全非均质性区域的地域规律性的变化过程而推导的。由于地理学家研究的许多现象受这种区位过程支配,所以函数 $\pi(x)$ 应当是适用于描述并

尽可能解释各种现象的区域分布问题的研究。

张克东 摘译自《Economic Geography》,
VOL.42, NO.2, 1966

李乃艳 刘妙龙校

1979)。还有一项发现表明,地图读者能够发现存在于无级地图之间的统计相关上的微小差别。因此,彼得森得出结论说,无级等值区域地图在表示专题信息方面是有效的。

穆勒在1979年用由传真静电绘图机产生的连续色调地图重复了彼得森的研究。这种传真机器共能产生200多种灰度值,且其中每一种都没有可观察到的纹理。穆勒曾预料彼得森所得到的结果已受到过去实验中采用的相交晕线模式带来的视觉干扰的影响。然而,穆勒的结论却支持了彼得森的断言,即在无级地图之间进行目视对照至少与在经量子化的地图间进行对照一样好。因此,可以得出结论说,在地图对照中,由相交线地图模式引起的视觉干扰是微不足道的。这一点很重要,因为穆勒所使用的传真机器目前还不是一种常见的输出装置,而能产生相交晕线的线划绘图机却十分普遍。

别的研究者正在重新考察采用无级地图的问题。布拉采尔和乌塔诺(1979)在假定读者能够理解无级地图的基础上,提出了一些设计策略,以帮助读者判读由无级地图表示的、公认为是更复杂的显示。他们强调,为了有助于地图传输,仔细斟酌图例设计十分必要。考克斯(1976)也考察了图例设计问题,他提出在分级圆圈图例中采用参考分级符号对信息传递可具有显著的影响。

看来,关于量子化问题的争论还远没有结束(道不森,1980;穆勒,1980)。不过,制图学家们似乎已开始承认以下做法的可行性,即采用阅读偏难的更复杂的地图,以更精确地表示分布现象。因而,那种时有听闻的关于要简化地图的告诫已不再为所有制图学家顺从地接受了(莫里森,1980)。

双变量地图和地图对照 地图对照的中心问题是找出两个或多个变量之间位置上的共同趋势。对照两幅地图最基本的方法是简单的目视比较;这种技术在地理分析中常常十分有效。然而,同样面对某两幅地图,不同的读者对面状上的相关程度看法不尽相同(恩维因,1981,第188页)。产生主观差异的一个尤为重要的原因在于分级的间断点数字,除非所对照的两种分布的间断点类似,否则很难进行有效的目视对照。采用无级地图就可以减轻这个问题的严重性。

目视对照法的另一个缺点是:为了证实分布之间视觉上的一致性,地图读者必须边对照,边记住他所看到的地图上的各部分,例如,得记住一本图集的不同页的内容。在视觉回忆方面,某些读者强一些,而另一些读者差一些。某些读者对制图化简的基本原理内行一些,而另一些读者则差一些。而如果我们将两种分布结合在同一幅地图上,则加在视觉上的负担便可减轻。

地图对照的各种方法在文献上有大量记载。其中绝大多数涉及到在不同的分布间为一些统计相关性制出图来。在对照问题中已经运用了地图相互之间的相关系数、回归残差、面状吻合系数、洛伦兹曲线、系数填入以及模差异模式等等(克利夫,1970;……)然而,一幅表示统计相关性的地图是否能合理地代表目视模式对照的结果还是个问题,虽然这些技术肯定是分析制图中的有用补充,但这种数根概括的一个明显后果是分布信息的损失,继而,对于联合形成地理相关性的各个分布的独立贡献,人们的认识程度便减少了。

关于使用和分析双变量地图的文献寥寥无几。其中多半是围绕美国人口调局在多色等值区域地图上双变量关系的光谱编码问题;不幸的是,人口调检局一般只关心生产这种

地图的技术可行性(巴拉巴, 1975; ……)。

已出版的关于采用双变量地图实验研究的报道很少。奥尔森(1981)在其中一篇中比较了地图用户在诸单幅地图之间及在以光谱编码的双变量地图上的各种分布之间估计相关性的能力。她发现有些时候用双变量地图,其相关性估计要好些。从整体看,奥尔森得出结论说:“实验提供了证据,表明正是关于具有类似值的区域相结合的信息能提供人们毫不困难的使用,而且一般说来,用户不会被这种地图过多的搞混,在看法上也不会受到消极的影响”(奥尔森,1981,第275页)。不过,奥尔森发现,受实验者们在对人口调查局所使用的色彩的相关性方面,几乎看法各异,莫衷一是,这些色彩来自人口调查局自己设计的带有明暗阶梯的正方形矩阵图例。这表明光谱序列的色彩不会使人们在头脑中建立起符号和数值之间在等级次序上的一致联系,这一点与分级的灰色调系列不同。所以,在单色的或纹理符号化系统范围内的模式变化倒能更好地服务于地图传输和地理分析。

程序BICHOR的特性 无级双变量等值区域地图可以看成是对相关性的统计度量的一种视觉模仿。另外,有许多迹象表明,这些地图是值得在专题传输中采用的。当然,这种断言还有待于我们用更严格的方式进行考验。传统的等值区域图方案有着分级间隔上的固有矛盾,看来地图读者对连续变化的多色符号的意义就存着识别上的困难。在这种进退两难的情况下,我们可以采用无级的相交晕线符号。

一个名叫BICHOR的、用BASIC语言写成的计算机程序已被研制出来,它可以在Tektronix CRT屏幕上或在数控笔式绘图机上产生无级双变量等值区域地图。BICHOR用依比例隔开的水平晕线来显示一个变量的地理模式,用依比例隔开的竖直晕线显示第二个变量的地理模式,如图1所示。这样,该程序把托布勒(1973)的单变量方案用在双变量的场合中。模式的视觉密度用线与线的间隔表示,而视觉密度与数据密度成正比例。适当的线划间隔是通过每个组数据的线性转换而取得的。该转换使每组数据都分布在0.0到1000之间。每种变量的范围或是扩大或是缩小,以适合普通的视觉比较。

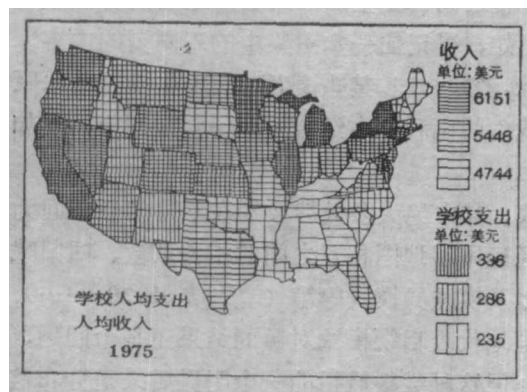


图1 一幅关于人均收入和学校支出的无级双变量地图。(数据来源:美国人口调查局)

是独立图例还不清楚,这一点有赖于进一步的实验分析。

图例参考分级值分别对应重新调整了比例的值域的25%、50%和75%。布拉采尔和

双变量地图图例的结构无论在专题或分析目的方面对地图用户的帮助都是十分重要的。常规方法已有将双变量符号组合表示在一个方阵中。然而,这类图例只是被用于量子化数据,而且往往分级过多(奥尔森,1981)。

显然,要在无级双变量地图上表示出所有可能的模式组合是不可能的。为了清晰,也为了强调出地图的双变量特性,在我们所研究的无级双变量地图中,为每个变量独立设置了图例。然而必须承认,最优的图例设计究竟是方阵还是

乌塔诺(1979)在单变量无级分级地图上使用了类似图的方法。还有一些人提出参考符号应该具有整个数值范围的跨度(张,1980)。这里争论的问题是选择适当数目的图例符号,以及在无级数据的连续统一体中选择具有代表性的间断点。正如图例结构的设计一样,这个问题也还没有解决。

无级双变量地图地应用和判读 穆埃尔卡评论道:“看来,当被比较的模式之一迭置到另一模式之上时,进行对空间相关性的视觉判断最为容易”(1978,第239页)。BICHOR允许这种迭置,也能被用来以平行线符号法建立无级单变量地图(见图2)。在地图对照中,对用户来说,既看分布的双变量,又看其单变量的构成部分是个可取的方法(奥尔森,1981)。

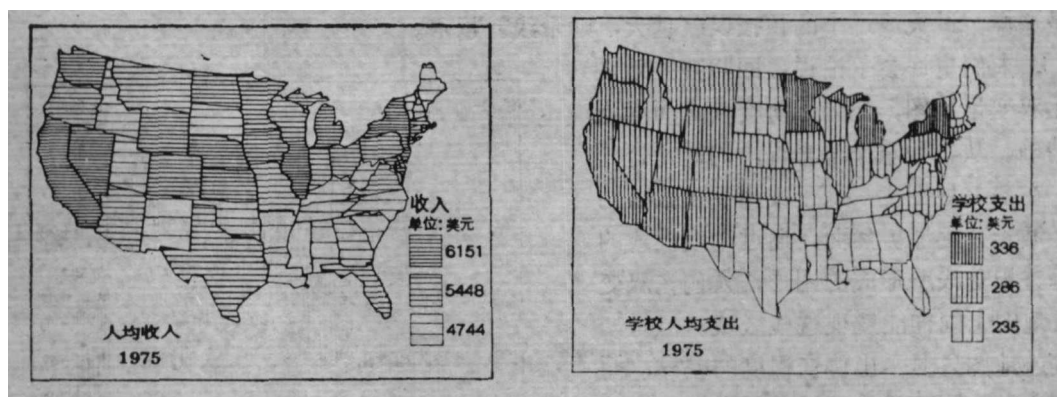


图2 无级单变量地图。(这两种分布与图1的分布是一样的。数据来源:美国人口调查局。)

在无级双变量地图上进行的两个变量之间的符号相关方式是独特的。它们比粗略的一瞥需要更多的时间,而且还需要作一点解释。不象在单变量地图上判读相交晕线那样简单,只把注意力放在这些地图的模式密度上,因为对变量相互之间关系的观察是模式方向所共同构成的函数。

图1可以用来表明双变量视觉相关的概念。自变量——人均收入,用水平线模式符号化,因变量——学校人均支出,用竖直线模式符号化。这些变量之间的相关程度部分地是由这两种线划模式垂直相交所形成的矩形的高与宽之比来表示的。这种比值越小,变量之间的吻合程度越高。因此,正如在科罗拉多州和华盛顿州所能看到的,由正方形构成的模式,就表示收入与学校支出之间一种紧密的相关关系。凡是具有水平方向的长方形的州,如德克萨斯州或北达科他州,就表示学校支出小于按收入计算的期望值。凡是具有竖直方向的长方形的州,如新墨西哥州或犹他州,表示学校支出超过了按收入水平的期望值。

不仅矩形的形状能反映变量之间的相对联系,而且矩形的大小也能反映一种更为绝对的关系。例如,缅因州和华盛顿州都有着一一种收入与学校支出之间的相对平衡关系,但是其正方形的大小截然不同。地图用户很容易得出结论:虽然每个州收入与学校支出在每个变量的值域内是平衡的,但是其中华盛顿州无论是收入还是学校支出都较高。

无级双变量地图最主要的用途是描绘出地理分布之间的协变性。地图用户也许很想

看到它在因果性方面的应用,但是有些变量说不清是自变量还是因变量。再说,因果性常常是循环的,因此,在一种情况形下的因变量,在另一种情形下可以独立起作用(托马斯,1968,第326页—349页)。例如,在图1中,人们可以说,人均收入为学校人均支出设立了一个上限,但从长远的观点看,公共学校的质量所含的结果性变量又影响到该州内的家庭经济状况。

图3是一幅表示回归残差的地图。这里,假定学校人均支出是人均收入的函数(虽然,如前所述,从长远观点看,后者也可以是前者的函数)。根据比尔森0.69的产品动差相关,我们能够推断出,州与州在收入上的差别,从统计上说,约占州与州在学校支出上差别的50%。用竖直线表示具有负的残差,那些学校开销超过按收入作出的预报的州就是这样,用水平线表示具有正的残差,那些学校开销低于按收入作出的预报的州就是这样。过大或过小的预报误差越大,则晕线就越密。

人们也许会争论说,回归残差地图比双变量地图,在显示协变性方面更为精确。从某种意义上来说,的确是这样,在残差地图上,低于或高于预报值的区域十分明显。然而,通过晕线模式的垂直相交线所形成的图形,人们在双变量地图上同样能发现这些区域。但是,残差地图只显示出协变程度的差异,而双变量地图可以分别显示两个变量的面状模式,而都撇开另一个变量。这样,例如在残差地图上显示出北达科他州和内华达州教育经费偏低的比例相当,而双变量地图则显示出内华达州,收入和

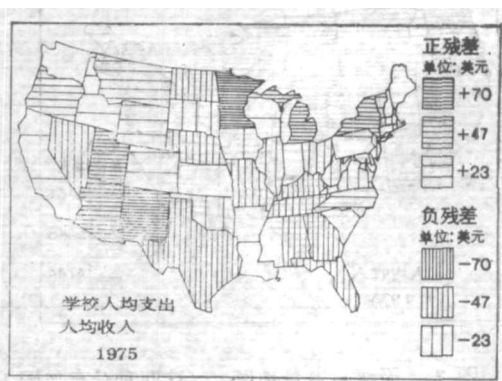


图3 关于收入(自变量)与学校支出(因变量)无级双变量回归残差地图。(比尔森的 $r=0.69$;回归方程是 $Y=12.32+0.05X$ 数据来源:美国人口普查局)

教育经费在二者之间较高。残差地图是双变量地图的一种有用的补充,反之亦然。然而,如果只用残差地图,将会使我们损失大量信息。何况,双变量地图在直观水平上易于被理解,而残差地图要求读者熟悉统计相关和回归。

在双变量地图的另一种用途中,人们可以从好几个角度来考察时间系列。使用一台能产生高分辨率矢量图的微型计算机,例如Tektronix 4052或4054,人们可以在一幅双变量图的迭盖透明片上显示相同变量不同时期的一系列组合。例如,可以用无级双变量式的“差异地图”来显示美国人口密度的变化趋势(克利夫,1970)。使用以某一时间段为参考的双变量地图时间系列也能用于异常情况的分析,其方式与回归残差地图几乎相同,但却没有那种方法固有的、大量损失视觉信息的弊端(托马斯,1968,第333页)。

结论 莫里森(1980)在最近关于计算机技术现在和将来可能对地图制图活动产生影响的评论中说,他发现有两种背道而驰的倾向。第一种是向更简单的地图发展。这种地图除显示地理信息系统的内容外,只需要注意一下比例尺或图名这些附属方面就行了。第二种是“相反,人们可以在数据之间进行更复杂的处理,这样可以把数据结合成对地图用户来说是更高级的一种抽象”(莫里森,1980,第21页)。

本文提出的作为一种分析工具的无级双变量地图代表属于上述第二种倾向的一个发展。它们不是简单的地图。不过，它们具备三个重要的性质。第一，它们消灭了量子化误差，这种误差在地图对照过程中可能是一种相乘关系而不仅仅是相加的。在常规等值区域地图上选择不当的分级间断点，如果不说是易于颠倒，也至少是易于掩盖面状协变性。第二，它们放低了由常规单变量的地图与地图对照所强加给人类的记忆上的要求。几乎没有一个地图读者在看一幅地图时，有本事“凭脑子的眼睛”来把握另一幅图上所有差异。第三，它们避免了多色双变量地图带来的问题，即读图者不能很好判读表示等级和间隔值变化的光谱波长。平行线之间的距离量度，对地图读者来说更易理解。

地图制图学家已倾向于把地图看成是在一处生产、而要给别处多半是未知的读者使用的一种产品。因此，地图设计的重点往往放在为相对不内行的地图读者削减地图信息上，以迎合尽可能多的读者。结果，地图的简化已成为一种规范，尤其在学术性地图上更甚。

我们感到，地图制图学家应考虑到在地图作者与地图用户之间存在的大量不同关系的可能性。随着易于做到的和真正的人机对话制图的出现，在某种意义上说，地图用户将变为制图者。因为地图制作更接近于一种有的放矢的工作，地图用户/制图者将能够根据各自使用信息的能力来增加，或减少视觉显示的复杂性。在我们看来，凡是在地图生产中喜欢标留信息而不是丢弃信息的人，应该采纳别的符号化方案，例如，无级双变量等值区域地图。

费立凡译自 《The American Cartographer》

第十一卷，1984年第1册。张天桢校