

平成22年 6月 3日現在

研究種目：若手研究（B）

研究期間：2008 ～ 2009

課題番号：20760400

研究課題名（和文） 危機的かつ不確実性を伴う空間情報の分析手法の開発

研究課題名（英文） Development of a methods for analysis of uncertain spatial data on critical events

研究代表者

片岡 裕介（KATAOKA YUSUKE）

東京大学・空間情報科学研究センター・助教

研究者番号：90451898

研究成果の概要（和文）：

本研究では、特に健康被害や犯罪といった住民の日常生活を脅かす危機的現象の把握のための、発生地点の空間的傾向に着目した分析手法の開発をおこなった。危機発生地点およびそれを近似する点にもとづく密度分布の空間的差異を定式化した。さらに、この両者の差異を最小化することにより、危機発生地点分布を少数の点分布へ置き換える方法の提案をおこない、サンプルデータを用いてその解法を明らかにした。

研究成果の概要（英文）：

A method for spatial data analysis focused on distributions of incidence points is developed to recognize the spatial trend in critical events. Two probability density functions are based on the distributions of the known points of the past and unknown points by kernel density estimation method, and the difference between these functions is formulated as integrated squared error. A method that approximates the distribution of known points by unknown ones and its solution using point data are proposed by minimizing the spatial difference between the two distributions.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008年度	1,000,000	300,000	1,300,000
2009年度	900,000	270,000	1,170,000
年度			
年度			
年度			
総計	1,900,000	570,000	2,470,000

研究分野：工学

科研費の分科・細目：建築学・都市計画、建築計画

キーワード：空間情報分析、解析・評価、地理情報システム、都市計画、最適化

1. 研究開始当初の背景

近年、GIS（地理情報システム）の普及もあいまって、地域統計データをはじめとした

都市や地域の空間情報の中には、整備状況が進展しているものも多い。ところが、健康危機や犯罪の発生に関わる情報は、大体的場合

において特殊であり、またプライバシーや地域のイメージに影響を及ぼす問題から、位置や場所に関連付けられた空間情報としての収集は、今後も困難であることが少なくない。さらに、収集可能な情報であっても、精度に限界があったり、あるいは重要な要素が欠落していたりといった、不確実な要素をもつ可能性が、通常のデータと比較しても高いと推察される。

点分布パターン分析については、これまでも計量地理学をはじめとした分野において様々な分析方法が提案され、現在では点データの精緻な分析が可能になっている。また、地理情報システム(GIS)の普及により、空間データを用いた分析を必要とする分野においても空間分析を行う環境が格段に向上している。なかでも、代表的な空間分布である点データ、および点分布に対する認識は高く、今後一層重要性が増すと考えられる。

2. 研究の目的

本研究では、都市における空間現象のなかでも、特に健康被害や犯罪といった住民の日常生活を脅かす危機的現象の把握のための、発生地点の空間的傾向に着目した分析手法を開発することを目的としている。

点パターン分析として頻繁に用いられるのが、点分布を確率密度で表される連続分布へと変換するカーネル密度推定法である。カーネル法では、各点が密度関数に置き換えられることで、その1点のみに情報を持たせることを避けることができ、例えば発生地点に不確実性がある際には、特に都合が良い。

本研究のように危機的現象を扱うにあたって、まず過去の状況を表す場合、そして現象の将来的な状況を表す場合の、2つの密度分布の関係性について検討する。同じく、このカーネル密度の比較に関わる問題とは、計量地理学や空間統計において重要な課題とされている。

カーネル法の推定量による点分布の比較を考えると、直感的には確率密度関数の差を考慮すれば良いことになる。従来の方法では、両者の信頼区間を得るために、各々の推定量の標準誤差を計算することとなる。これに対して本研究では、より自然な確率密度関数間の差の表現として、その積分二乗誤差を用いることとし、これをもって現象の過去と将来における空間的な関係性を議論する。

そこで本研究では、過去の観測されたデータ点の分布があるときに、これを将来のある一定期間内を表す、別の点分布に置き換えることによって、危機的現象の発生の空間分布や、影響の地理的差異に関する分析を行うことを目的とする。

3. 研究の方法

空間分析方法を開発するにあたり、点分布における空間パターンについて考察する必要がある。そこで、まず実際に危機的イベントが発生した地点の分布パターンについては、過去に数多く発生した地域が、潜在的に危険性が高い地域であると考えられる。すると各発生地点周辺の危険度は、発生地点で最大で周囲に拡がるにつれて減衰する確率密度分布で表現することができる。また、本研究では、同地点で現象が再発する可能性や位置誤差などを考慮した場合、与えられた点分布を連続的な密度分布として扱うことが適切であると考えている。そこで、危機現象の過去の一定期間の発生地点を、カーネル密度推定法を用いて密度分布へ変換する。

一方で、この発生地点分布を置き換えようとする未知の点分布についても、同様に各地点を最大とする確率密度分布で表現できる。以上より、両者に関する地域全体の確率密度分布についてはカーネル法を用いて推定できる。なおカーネル法とは、点分布を形成する各点の中心に、カーネルと呼ばれる密度分布の山を置き、それら全てのカーネルの合計をもって全体の確率密度関数を求める密度推定法である(図1)。

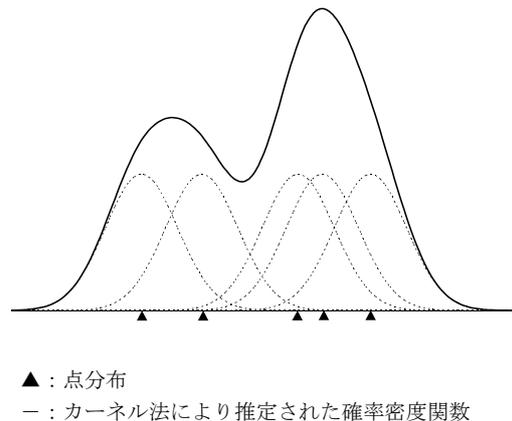


図1 カーネル密度推定法 (1次元上)

次に、過去の発生状況を表す密度分布を、限られた少数の点分布からなる未知の密度分布で近似することを考える。つまり、この未知の密度分布こそが、過去の空間的傾向から導かれた、将来的に予測される現象の発生状況とみなせる。その具体的な手続きとしては、元の点分布による確率密度関数と、未知の点分布による確率密度関数との差を定式化し、この積分二乗誤差を計量する。言い換えると、任意の点における発生地点の密度分布の差の領域全体の総和が、過去と将来の現象発生に関する空間分布間の差異を表すことになる(図2)。

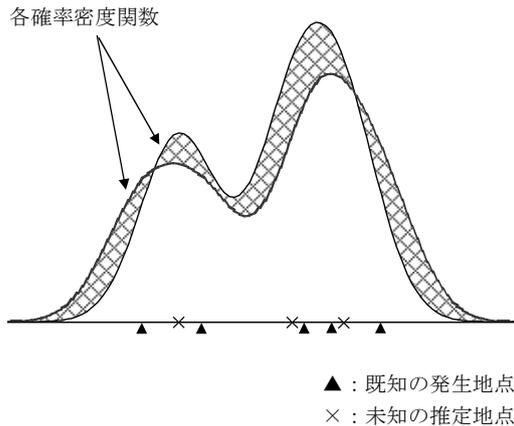


図2 異なる確率密度関数間の差異(1次元上)

カーネル法により推定された、2つの連続的な密度関数の積分二乗誤差の定式化に伴う積分計算においては、データ分析における計算量の軽減のために不定積分の計算をおこなう。積分計算においては、数値積分などの数値計算法についても計算し、不定積分を用いた場合との結果の比較をおこなう。また、点分布を連続分布に置き換える際に用いるカーネル関数について、単純な関数型から一般的な正規分布を用いた場合について、計算可能性を検討する。

この積分二乗誤差を目的関数とすると、将来の発生地点となる各位置座標に関する多変数最小化問題を定式化することができ、これらの解を求めることによって、長い過去の状況に対して、一定期間に起こり得る現象の発生地点が推測できる。また、最適化計算における計算効率の評価によって、アルゴリズムの改良などをおこない、分析方法としてより汎用的で有用なものとする。加えて、未知の点分布の個数と効率性の関係など、今後の課題をふまえた研究の応用性について吟味をおこなう。

さらに、実際にサンプルデータを用いながら、得られる結果の検証をおこなう。モデル化された地図上に結果を図示するなどして、既存の地図上での視覚化を想定した、効果的な表現法を検討する。

最後に、本研究の結果による地域情報地図の作製といった、研究成果の有効な活用例を提示する。

4. 研究成果

都市空間において生じる、危機的現象の過去の一定期間の発生地点を、カーネル密度推定法を用いて密度分布へ変換し、その過去の発生状況を表す密度分布を、限られた少数の点分布からなる未知の密度分布で近似するという方法を本研究において提案し、その解

法について明らかにした。

研究成果の具体的内容について、以下に示す。

(1) 2つの異なる点分布間の空間的差異の定式化、および差異の最小化による点分布近似法の提案

まず元の点分布による確率密度関数と、未知の点分布による確率密度関数との差異を表す、積分二乗誤差の定式化をおこなった。この定式化のための積分計算においては、カーネル関数として比較的単純な関数を用いた場合に限らず、一般的な正規分布を用いた場合であっても、その後のデータ分析において計算量を軽減できる関数が、積分計算の結果として得られることが確認された。また、対象地域を無限領域に設定できる場合において、さらに簡素化された数式で、2つの異なる確率密度関数間の空間的差異を記述することが可能であることが示された。つまり、地域の中心部に人口や活動量が集中する傾向がある一方で、周縁部においてはイベントがほぼ発生しないという地域条件においては、特に本方法が有効であるといえる。

続いて、この積分二乗誤差を目的関数とすることで、将来の発生地点となる各位置座標に関する多変数最小化問題を設定し、現象の将来的な発生地点を推測する方法論を示した。

(2) 点分布近似法の解法の検討と地域空間における危機的現象への応用化

実際の危機発生現象を表す空間データを用いて分析をおこなうことを想定した解法について示した。計算を行う最適化アルゴリズムには、逐次二次計画法 (Sequential Quadratic Programming) を用いた。この方法は共役方向を求めることで、2次収束性を有する手法となっている。この方法を本研究における最適化計算に用いることで、多変数の最適化を決定論的に行うことができるとともに、未知パラメータの数が少なく、最適解の範囲が大体予想がつき、かつパラメータ感度が鈍い場合においてこの方法は効果を発揮する。しかしその一方で、局所解に陥りやすいことと探索できる空間が狭いことが欠点であるため、適切な探索空間の範囲を指定することが求められる。本研究においても、特に推定する地点が多い場合においては、地域内の危機発生密度分布の様態を注視した上で、解の初期値を必要に応じて試行錯誤的に決定する必要がある。

続いて、実際の都市や地域に発生する現象に対する本方法の応用化を想定し、現実の地域空間における分析の有効性の観点から検証をおこなった。サンプルデータを用いた計算をおこない、分析方法としてより汎用的なものとした。その結果、過去の発生地点の密度分布と、それを近似する少数の地点の密度

分布との積分二乗誤差の最小化問題において、数十のオーダーの未知の各地点の座標値について収束解が得られることがわかった。これにより、点分布を正規分布のような連続関数を用いたカーネル法による密度分布であっても、計算実行の面から分析が可能であることが示された。ただし、当然のことながら推測する点が増加するにつれて、より初期値に依存する局所解となる傾向があるため、元の点分布の状況をふまえた分析が要求される。また、推測可能な地点が数十というオーダーで十分であるかという議論があるが、そもそも少数の地点で近似するという意義から考えると、元の発生地点はその何倍、何十倍の数になるであろうことから、対象とする空間範囲にもよるとはいえ、都市や地域に発生し得るイベントの多くが適用範囲内にあると判断できる。

ここで、(1)における両者の密度関数間の関係性に関する議論では、バンド幅は所与のものとして扱われているが、実際のデータを用いた際に、カーネル密度推定法においてはバンド幅選択の問題がある。バンド幅によって推定される確率密度が異なることから、その選択は慎重におこなう必要がある。バンド幅の選択法として、方程式解によるプラグイン法が多方面で有効であると言われており、本研究で既に与えられている発生地点の密度分布については、プラグイン法によるバンド幅を用いたカーネル法で推定することは適切であると考え、実際に実行可能である。しかし、これを置き換える未知の点分布を得る際の、座標値に関する多変数最適化計算において、変数である座標値を用いるプラグイン法の計算を同時におこなうことは、少なくとも有意義な解を得るという意味でも、現時点では実行が困難であると判断した。そこで本研究においては、未知の点分布にもとづく確率密度関数を推定する際のバンド幅は、容易に導出が可能である「経験則」の方法により求めるものとした。

(3) 地域現象における空間パターンの視覚的表現

最後に、分析結果の表現法として空間パターンの視覚的表現の検討をおこない、危機情報の伝達において、点密度分布の等高線や濃淡図による表示が有効であることを確認し、本研究成果の活用可能性を吟味した。また、本研究で提案される分析方法を適用する、危機現象の候補の一つとして、感染症流行などの健康危機現象を採り上げ、各地点の流行発生状況を表す点データをもとづいた密度分布の傾向について吟味するなど、本研究の応用化を試みた。

なお図3は、サンプルデータとしてバラツキが比較的大きい、危機現象の発生地点を用いて、それらを近似する20地点を推定した

結果となっている。図によれば、発生地点が集中している中心部、そして中心部の他に局所的に集中している場所の幾分中心側に、推定地点が得られていることが確認できる。また、図中の曲線は、過去の発生地点、未知の推定地点のそれぞれにもとづく確率密度の程度を示す等値線である。同じくこの結果からは、推定地点が20地点という少ない数ということもあるが、(2)で述べたように「経験則」によるバンド幅を用いたカーネル法により、推定された点分布にもとづく密度分布が、真の密度分布よりも過度に平滑化されている可能性が疑われる結果となっている。

この等値線を併用した結果の表現により、例えば地図を背景とした場所の安全度の表示なども可能であり、地域情報地図の作製などへの活用についても十分期待できる。

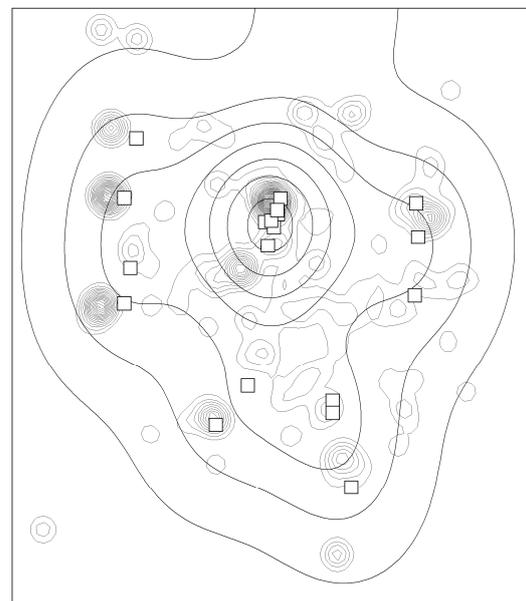


図3 過去の発生地点の近似 (20点)
 (□: 未知の発生地点、○: 未知の推定地点の密度分布、●: 過去の発生地点の密度分布)

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計1件)

① 郡山一明, 片岡裕介, 竹中ゆかり, 浅見泰司, 高橋邦彦, 丹後俊郎, 健康危機管理と小学校欠席状況サーベイランス, 保健医療科学 57(2), 2008, pp. 130-136, 査読有

〔学会発表〕(計1件)

① 片岡裕介, 健康問題に関わる空間情報分析, CSIS SYMPOSIUM 2009, 2009年12月18日, 東京大学駒場リサーチキャンパス

6. 研究組織

(1) 研究代表者

片岡 裕介 (KATAOKA YUSUKE)

東京大学・空間情報科学研究センター・助
教

研究者番号：90451898