

都市間非対称流動データの視覚化

大保 和子, 善家 大輔, 山本 芳嗣

都市間の非対称流動データとは、同一都市内での流動量が都市規模によって大きく異なり、また、ある都市から他のある都市へ流動量がその逆方向の流動量とは異なるようなデータのことである。都市を3次元空間に布置することによってこのようなデータの構造を視覚化する方法をいくつか提案する。データとして茨城県85市町村間の通勤通学者数を取り上げ、提案した方法によって求めた市町村の3次元布置を報告する。また、得られる布置が市町村の地理座標を反映するように加えたいくつかの工夫と試行錯誤にも触れる。

キーワード：非対称流動データ、多次元尺度法、重力モデル、通勤通学者数データ

1. はじめに

表1は、平成12年国勢調査従業地・通学地集計の「常住地による従業・通学市区町村、男女別15歳以上就業者数及び15歳以上通学者数」[1]をまとめ直した茨城県85市町村（当時、平成18年3月末には44市町村となる予定）間の通勤通学者数データの一部である。例えば「水戸」の行には、水戸から水戸への内々通勤通学者数と、水戸から他の市町村（以降は一括して町と呼ぶ）への通勤通学者数が書かれている。県央の中心地である水戸から南へ35キロほどの石岡への通勤通学者数516人に対して、石岡から水戸へはそれに倍する1,274人の通勤通学者がいることが分かる（表1の下線部と図1）。このようなデータの様子が一目で見て取れるような「地図」を描きたいと思って行った試行錯誤をご報告したい。

2. データの特徴

以降、表1のデータを $C=[c_{ij}]_{i,j=1,2,\dots,n}$ で表すが、このデータには以下の特徴がある。

a) 親近度： c_{ij} や c_{ji} が大きい町 i, j 同士は関係が

表1 茨城県85市町村間の通勤通学者数

	水戸	日立	土浦	石岡	...
水戸	107,045	2,422	1,113	<u>516</u>	
日立	4,016	84,689	108	46	
土浦	1,091	104	45,942	787	
石岡	<u>1,274</u>	86	2,863	5,938	

おおほ かずこ, ぜんけ だいすけ, やまもと よしつぐ
筑波大学 大学院システム情報工学研究科
〒305-8573 つくば市天王台1-1-1

深い。

b) 非ゼロ対角要素：通常 $c_{ii} \neq 0$ である。

c) 非対称性：通常 $c_{ij} \neq c_{ji}$ である。

このようなデータから相互に通勤通学者数の多い町は近くにあるように地図を描こうとすると、多次元尺度法が選択肢として浮かんでくる。 n 個の対象間の距離とおぼしきデータ $D=[d_{ij}]_{i,j=1,2,\dots,n}$ が与えられたとき、すべての i, j について

$$d_{ij} \approx \|x^i - x^j\| \quad (1)$$

となるように、適当な次元の空間の点 x^1, x^2, \dots, x^n を決めるのが多次元尺度法であり、点 x^i を対象 i の布置という。距離とおぼしきデータと書いたのは、少なくとも対称性 ($d_{ij} = d_{ji}$)、非負性 ($d_{ij} \geq 0$) と対角要素がゼロ ($d_{ii} = 0$) くらいは成り立っているとの気

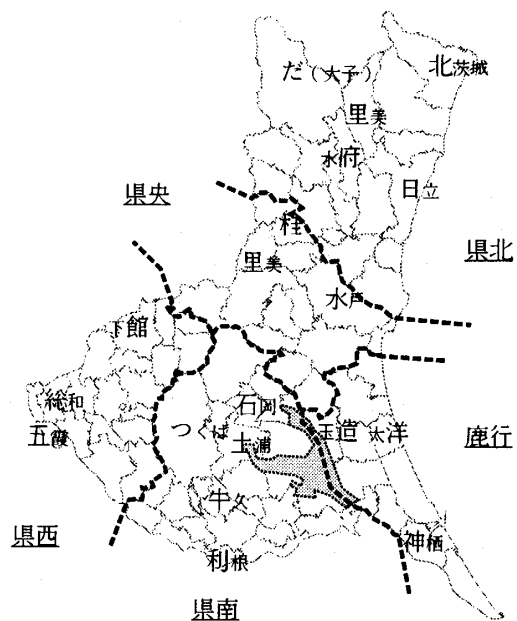


図1 茨城県の市町村と県総合計画にある地域境界、鹿行と県南との境界にあるのは霞ヶ浦

持ちである。表1のデータ C の対角要素 c_{ii} がすべて同じ値、例えば Δ で、どの非対角要素よりも大きい場合には、 $d_{ij}:=\Delta-c_{ij}$ によって $D=[d_{ij}]$ を作れば、 D は非負性と対角要素がゼロであるとの条件を満たすことができるが、対称性は成り立たない。しかも、表1の C の対角要素は各町の人口を反映して大きく異なっている（このような点を考慮した多次元尺度法が研究されている。千野[2]を参照）。さらに、図1に示したように町は実際の地理上の2次元座標を持っており、得られた布置がこの座標を反映するようにしたいし、目に見える地図を作りたいのであるから、布置の次元は高々3次元止まりにしたい。

3. モデル

ここでは、データ C から市町村の親近度と非対称性を見るために3次元に布置することにし、町 i の布置を $(x^i, y^i) \in R^2 \times R$ で表す。ここで、 $x^i=(x_1^i, x_2^i)^T$ は町 i の布置の初めの2次元座標からなるベクトルである。町 i と町 j の相互の通勤通学者数 c_{ij} や c_{ji} が多いほど、その布置 x^i と x^j を平面上の近くに置いて、両町の親近度が高いことを表現したい。また、町 i から町 j への通勤通学者数が逆方向の通勤通学者数を上回っている場合、すなわち $c_{ij} > c_{ji}$ の場合には、町 i の標高を町 j の標高よりも高くして、非対称性を表現したい。もちろん3つの町 i, j, k の間の $c_{ij} > c_{ji}$, $c_{jk} > c_{kj}$ かつ $c_{ki} > c_{ik}$ という関係も考えるが、このような循環構造は、対象としている通勤通学者データではあまり多くはないだろうと考えた。

3.1 C を分解するモデル

データ C は

$$C = \frac{1}{2}(C + C^T) + \frac{1}{2}(C - C^T)$$

のように対称行列 $S=[s_{ij}]=\frac{1}{2}(C+C^T)$ と歪対称行列 $T=[t_{ij}]=\frac{1}{2}(C-C^T)$ との和に分解できる。 S から $x^i=(x_1^i, x_2^i)^T$ を、 T から y^i を作るのが初めのモデルである。 x^i については、

i) 単調性: $\sigma(a_1, a_2, \beta)$ は、 a_1 と a_2 の増加関数で、 β の減少関数

ii) 対称性: $\sigma(a_1, a_2, \beta) = \sigma(a_2, a_1, \beta)$

なる性質を持った関数 $\sigma: R^3 \rightarrow R$ を用意し

$$s_{ij} = \sigma(s_{ii}, s_{jj}, \|x^i - x^j\|) + \varepsilon_{ij}^s (i \neq j \text{ について}) \quad (2)$$

として、誤差項 ε_{ij}^s の2乗和を最小化する。つまり、

$$\text{minimize } \sum_{i=1}^n \sum_{j=1; j \neq i}^n (\varepsilon_{ij}^s)^2 \text{ subject to (2)}$$

を解いて布置 x^i ($i=1, \dots, n$) を求める。ここでは関数 σ として重力モデルに似た

$$\sigma(s_{ii}, s_{jj}, \|x^i - x^j\|) = \frac{\sqrt{s_{ii}s_{jj}}}{\|x^i - x^j\|^2}$$

を用いた。また、

i) 単調性: $\tau(\gamma_1, \gamma_2)$ は γ_1 の増加関数で、 γ_2 の減少関数

ii) 歪対称性: $\tau(\gamma_1, \gamma_2) = -\tau(\gamma_2, \gamma_1)$

なる関数 $\tau: R^2 \rightarrow R$ を用意し

$$t_{ij} = \tau(y_i, y_j) + \varepsilon_{ij}^t (i \neq j \text{ について}) \quad (3)$$

の誤差項 ε_{ij}^t の2乗和を最小化する。つまり、

$$\text{minimize } \sum_{i=1}^n \sum_{j=1; j \neq i}^n (\varepsilon_{ij}^t)^2 \text{ subject to (3)}$$

を解いて布置 y^i ($i=1, \dots, n$) を求める。ここでは

$$\tau(\gamma_1, \gamma_2) = \gamma_1 - \gamma_2$$

とした。なお、Plane[3]は米国の州間人口移動を説明するために、州間距離を補正する際に類似の重力モデルを用いている。

3.2 C を分解しないモデル

このモデルは、 C を分解しないで、まとめて

$$c_{ij} = \sigma(c_{ii}, c_{jj}, \|x^i - x^j\|) + \tau(y_i, y_j) + \varepsilon_{ij} \quad (4) \quad (i \neq j \text{ について})$$

として、前と同様に

$$\text{minimize } \sum_{i=1}^n \sum_{j=1; j \neq i}^n (\varepsilon_{ij})^2 \text{ subject to (4)}$$

から布置 (x^i, y^i) ($i=1, \dots, n$) を求めるものである。

3.3 試算した結果

両モデルの結果にほとんど差がなかったため、以降は C を分解するモデルによる結果を示しながら、主に望ましい2次元の布置 x^i を得るための試行錯誤を報告することにし、3次元目の布置 y^i は最後の結果

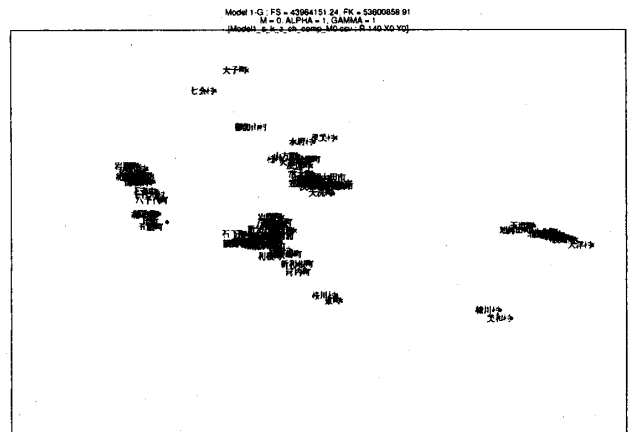


図2 モデル(2)から得られた2次元布置

についてのみ示すことにする。図2は(2)で定義されるモデルから得られた布置 x^i である。七会村など市町村合併で近ごろ消えた地名が見えること、地名が小さいことはご容赦願いたい。県北、県央、県南、県西、鹿行の地域にある町々が塊を作っている様子がなんとなく窺えるが、県の形は失われている。

4. 地理座標を反映させるために

前節で得られた布置は、実際の町の地理座標からずれているが、そのずれの大きさや方向を把握することが難しい。実際の地理座標からのずれが見えてこそ、その理由を通勤通学者数に求めることができるのである。そこで、町の地理座標をある程度維持するモデルをいくつか考えた。

4.1 変位の制限

町 i の地理座標を $X^i = (X_1^i, X_2^i)^T$ で表す。まず考えたモデルは地理座標と布置との差異をペナルティに組み込んだ

$$\text{minimize } \sum_{i=1}^n \sum_{j=1, j \neq i}^n (\epsilon_{ij}^s)^2 + M \sum_{i=1}^n \|x^i - X^i\|$$

で、ペナルティパラメータ M を徐々に大きくして地理座標からのずれを抑えるものである。 M を大きくすると徐々に茨城県の形が現れてきたが、紙面の都合で詳細は割愛する。

4.2 方位の維持

町相互の東西南北の位置関係を布置に引き継がせようとしたのが次のモデルである。例えば、町 i が町 j の北東にある場合、つまり、 $X_1^i \geq X_1^j$, $X_2^i \geq X_2^j$ である場合、両者の布置 $x^i = (x_1^i, x_2^i)^T$, $x^j = (x_1^j, x_2^j)^T$ に対して同様の方位の関係を要請する制約 $x_1^i \geq x_1^j$, $x_2^i \geq x_2^j$ を追加して解いた。同様の制約の作り方を古藤[4]は時間地図を描くのに用いている。結果を図3に

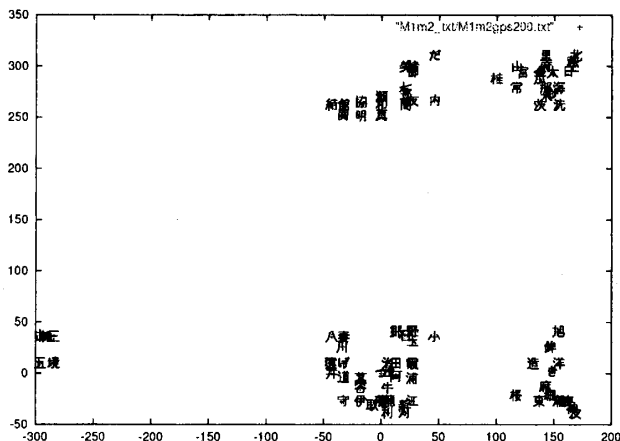


図3 方位を維持した布置

示す。おおよそ県北、県央、県南、県西、鹿行の地域の町々が固まっている点は図2と変わらないが、どの町もそれより北(東)にある町を越えて北(東)側に布置を持つことができないため、町々が碁盤の目で区切られた矩形の中に塊となって押し込められているようで、自然さに欠ける。

4.3 回転の維持

次に考えるのは回転である。ある町 i に視点を置いて、反時計回りに目を転じるに従って町 j , 続いて町 k が視界に入ったとする。このとき、各町の布置についても同じ関係を維持するのがここで取り入れた制約である。町 i の地理座標 $X^i \in R^2$ を使えばこの制約は

$$\det[X^j - X^i, X^k - X^i] \det[x^j - x^i, x^k - x^i] \geq 0 \quad (5)$$

と書ける。町 j に視点を置いた際の町 i と k の見える順序についての制約も、町 k に視点を置いた際の町 i と j の見える順序についての制約もこの1本の式で表現されているので、町の3つ組について1本の制約を作れば良いことになる。それでも組み合わせが多い。85の町から3つの町を選ぶと、10万に近い本数の制約ができてしまう。ちなみに図2に示した布置はこのうちの約4万本の制約を満たしていなかった。何らかの方法で、全体として茨城県の形を彷彿とさせるような布置が得られる少ない制約を抽出する必要がある。

まず、図2の布置が満たしていなかった約4万本の制約から無作為に4千本を取り出して制約として組み込んでみたが、得られた布置は取り込まれた制約によって大きく異なっており、あまり芳しいものではなかった。

そこで、水戸を上記の制約(5)の町 i とし、水戸以外の町を j , k とし制約(5)を追加してみた。水戸を原点に布置した結果を図4に示す。回転を維持する制約

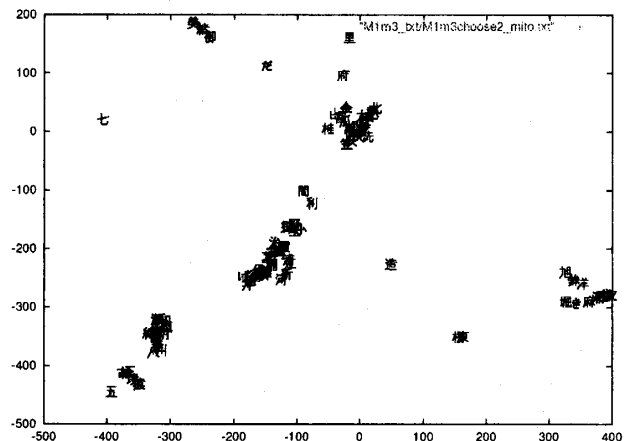


図4 水戸を中心に回転を維持した布置

がバインディングな制約となり，明らかに水戸を中心にして放射状に町が布置されており，不自然である。

遠く離れた町の間での回転は維持されなくても良いが，近い町の間については維持したい。そこで，85市町村を母点としてドロネ三角形分割¹を作り，三角形の頂点になっている3つ組について回転を維持する制約を入れることとした。同様の制約の作り方は古藤[4]にも見ることができる。図5にドロネ分割を示す。ここには163個の三角形ができています。この163個の三角形から作られる163本の制約を追加して解いた結果が図6である。茨城の形を維持しつつ通勤通学の多い町々が塊を作っているのが見て取れる。なお，この布置では10万近い町の3つ組の中の16,100組が回転を維持できていなかった。

4.4 霞ヶ浦があった

前節の図6は茨城県の形と町々の親近度をよく再現

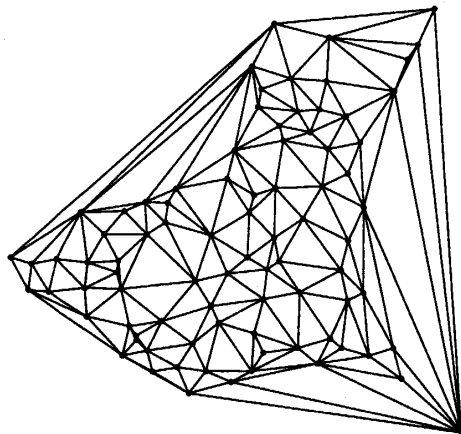


図5 85市町村のドロネ分割

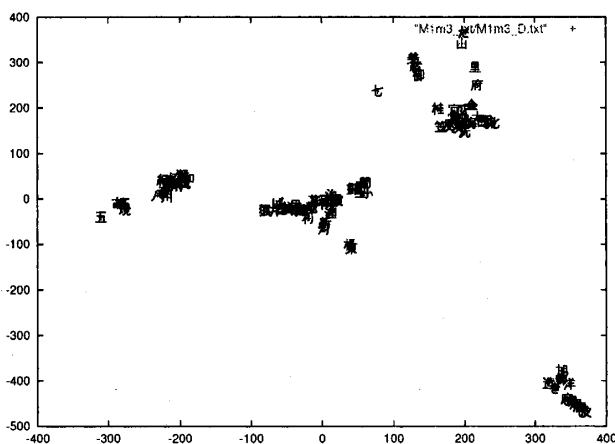


図6 ドロネ分割から得られた回転を維持する制約を付加した布置

¹ 与えられた3つの母点を周上に持つ円の中に他の母点が入っていないときに，この3つの母点を辺で結ぶ三角形から作られた無向グラフ。

しているが，そもそもの地理座標からの変化を見ることができていない。そこで，回転を維持する制約から外して，ペナルティ項として目的関数に組み込み，ペナルティパラメータ M を動かして得られる布置の変化を見ることとした。目的関数は

$$\begin{aligned} & \text{minimize } \sum_{i=1}^n \sum_{j=1; j \neq i}^n (\epsilon_{ij}^s)^2 \\ & - M \sum_* \det[X^j - X^i, X^k - X^i] \\ & \det[x^j - x^i, x^k - x^i] \end{aligned}$$

である。ペナルティ項の \sum_* は，ドロネ分割の三角形の頂点となった町163組と，前節の結果で回転が維持されなかった16,100組の計16,263組について合計を取ったことを示している。この結果を示す図7では，ほとんどの町が右上の塊につぶれ，一人大きく飛び離れて左隅に七会村が見える。右上の塊を拡大してみても町の地理座標は維持されていなかった。七会村を含む3つ組の町のペナルティ項が非常に大きな値をとることで目的関数値を大きく下げたため，このような布置ができてしまったと考えられる。これはペナルティ項が良くないせいである。

そこでペナルティ項の行列式の積 $\det[\cdot, \cdot] \det[\cdot, \cdot]$ を

$$\begin{aligned} & \min\{0, \det[X^j - X^i, X^k - X^i] \\ & \det[x^j - x^i, x^k - x^i]\} \end{aligned}$$

に置き換えた。以下は $M=0.05, 0.01$ として得られた図である。図8 ($M=0.05$) では茨城県がほぼ再現されている。図8で右下に布置されていた鹿行地域の町が，図9 ($M=0.01$) では水戸に向かって北に移動したことが見える。鹿行地域は，土浦・つくばを中心とする県南地域に近いが，間に霞ヶ浦を挟んでいるため，むしろ水戸周辺の県央地域との間に通勤通学者が多い。このことが，ペナルティパラメータ M を小さ

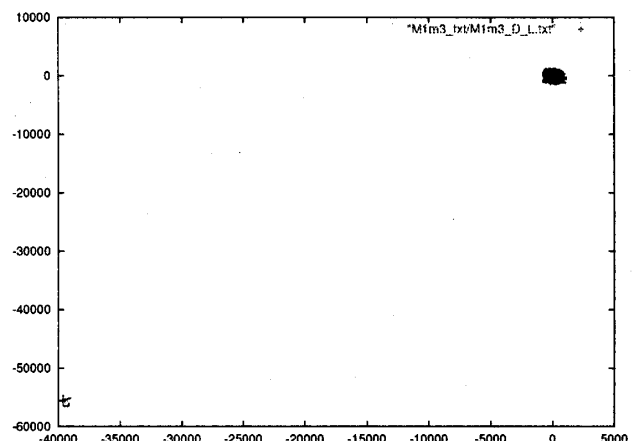


図7 ペナルティモデルによる布置 ($M=0.00001$)

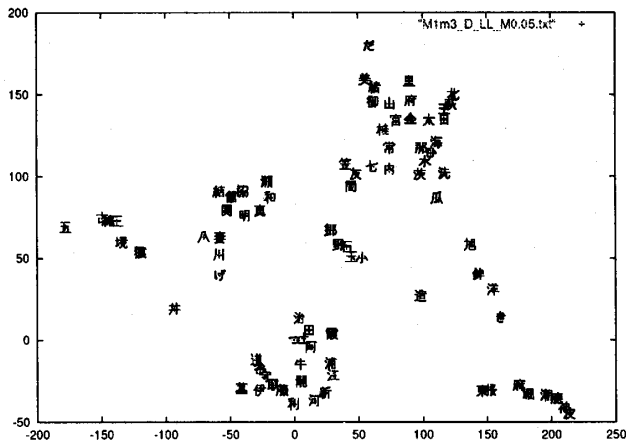


図8 修正ペナルティモデルによる布置 ($M=0.05$)

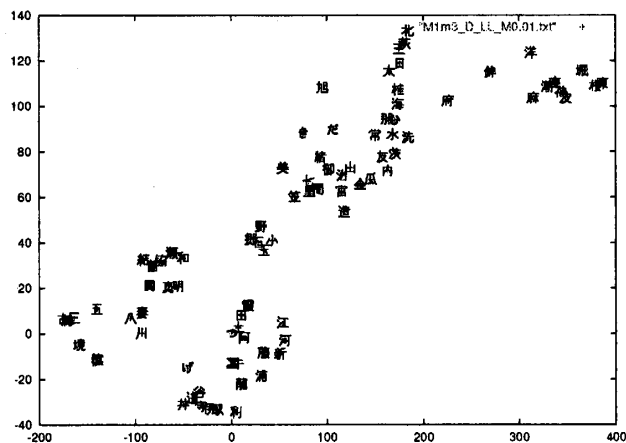


図9 修正ペナルティモデルによる布置 ($M=0.01$)

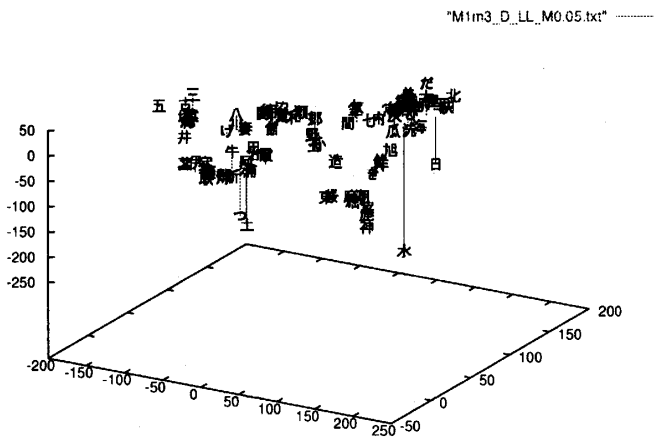


図10 修正ペナルティモデルによる3次元布置 ($M=0.05$)

くするに従って布置に反映された結果であると考えられる。

5. 3次元の布置と内々通勤通学者数の推定

初めに非対称な通勤通学者数データを表現する3次元の地図を描きたいと書いておきながら、これまで3

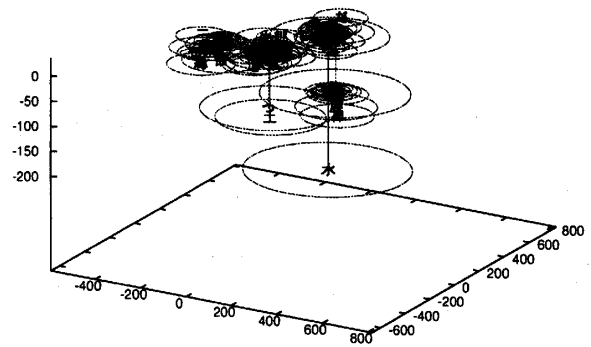


図11 内々通勤通学者人数も表現するモデルによる布置

次元目を描いてこなかったため、図8に対応する3次元地図を図10に載せる。町*i*の名称を示す文字が平面座標 x^i 、標高 y^i の位置に書かれ、そこから標高ゼロの点まで鉛直に線が引かれている。町の標高の平均がゼロになるように設定してあるので、負の標高を持っている町に向かう通勤通学者が多いことになる。目立って負の標高を持っている町は、県北の日立、県央の水戸、県南の土浦とつくばである。水の流れのごとく、高い町から近くに布置された低い町へ通勤通学があることがこの図から読み取れることを期待している。

以上説明してきたモデルはいずれも内々の通勤通学者数 $C_{ij}=S_{ij}$ を推定すべき対象としていない。この点を考慮するため、内々も含めた通勤通学者数の推定値を表す変数 z_{ij} ($i, j=1, 2, \dots, n$) を用意し

$$\begin{aligned} &\text{minimize } \sum_{i=1}^n \sum_{j=1}^n (S_{ij} - z_{ij})^2 \\ &\text{subject to } z_{ij} = \sigma(z_{ii}, z_{jj}, \|x^i - x^j\|) \\ &\quad (i \neq j \text{ について}) \end{aligned}$$

というモデルを設定した。図11は、ドロネ分割から得られた回転を維持する制約を追加して布置を求めた結果である。町の名前の文字を中心に水平に描かれている円の半径は $\sqrt{z_{ii}}$ であり、この円の面積によって内々通勤通学者の多寡を表そうとしている。見栄えの良い図にするためにはさらに工夫が必要である。

6. 補足

本稿中で解いた問題はいずれも非線形最適化問題であり、凸性すらない問題が大半である。また、2次元の布置 x^i を求める問題の目的関数値は、距離を変えない変換に対して不変であるため、最適解が無数に現れる。この問題点から少しでも逃げるために、2次元の布置を求める際にはつくば（ただし図4では水戸）の布置が原点になるように制約を加えたが、回転を抑えてはいない。また、3次元目の布置を求める際には、

布置の平均がゼロとなるように制約を追加している。

また、初めに触れた循環構造については以下の検証を行った。Cを分解するモデルで求めた布置(\hat{x}^i, \hat{y}^i)から誤差項を $\varepsilon_{ij} = c_{ij} - \sigma(s_{ii}, s_{jj}, \|\hat{x}^i - \hat{x}^j\|) - \tau(\hat{y}^i, \hat{y}^j)$ で求め、 $[\varepsilon_{ij}]$ をデータとして、そこに循環構造が残っているかどうかを、千野[2]の方法を使って布置を求めて確かめたが、循環構造はないといっても良い布置であった。

なお、以上の最適化問題は数理システムの数理計画法パッケージ NUOPT[5]を利用して解いたが、初期点の選び方などで多くの試行錯誤が必要であったことを付け加えておく。

謝辞 県総合計画について、筑波大学システム情報工学研究科の大澤義明先生に、茨城県の地図は、同じく渡辺俊先生にお世話になった。また、この試みを始めた頃、静岡大学の関谷和之、安藤和敏先生に相談に乗っていただいたし、東北芸術工科大学の古藤浩先生に

は初期の原稿にコメントをいただいた。最後に筑波大学システム情報工学研究科のプロジェクト「地理情報の解析と視覚化」の支援を受けたことに謝意を表したい。

参考文献

- [1] <http://www.stat.go.jp/data/kokusei/2000/jutsu1/08/index.htm>
- [2] 千野直仁『非対称多次元尺度法』現代科学社, 1997.
- [3] D. A. Plane, "Migration space: doubly constrained gravity model mapping of relative interstate separation", *Annals of the Association of American Geographers*, 74 (2) (1984) 244-256.
- [4] 古藤浩「いくつかの新しい時間地図とその特徴」日本オペレーションズ・リサーチ学会秋季研究発表会アブストラクト集 (2000) 72-73.
- [5] 株式会社数理システム <http://www.msi.co.jp/>