

農業利用における小規模風エネルギーの メッシュデータによる評価方法の研究

第2報 地形分類メッシュデータによる多次元解析法*

橋口 渉子**

要 旨

前報で予備解析を行った風の基礎データ、すなわち各月の各時間帯から成る12変数および対象88地点における風速2 m/s以上の頻度を目的変数とした。これに対する説明変数としての、地形分類メッシュデータの縮約および変換手法を呈示する。地点別原データは、該当地点をふくむメッシュを中心とする41×41個の3次メッシュであるが、このうち中央から放射状に16方向3列のデータだけを用いる。ほぼ9個のデータをふくむセルを作り、セル単位の8×13個のデータに縮約した。このデータから、地形パターンを表す説明変数を得るため、地点間類似性行列Sを作成し、Sを手がかりに各地点に数量を与える手法を示した。この手法は、説明変数のデータ行列は未知であるがSは既知としたときの回帰主成分の解の導出法に特色がある。内陸立地20地点のデータを用いて、この手法の数値例を示した。

1. はじめに

第1報で、季節、時間帯別の風エネルギー供給を表す基礎データとして、1、4、7、10月の各月、および各時間帯として1～8時(アサ)、9～16時(ヒル)、17～24時(ヨル)の各組み合わせによる12変数について、風速2 m/s以上の平均頻度を、対象88地点について算出した。そして、 i 地点、 j 月、 k 時間帯の頻度 y_{ijk} に対し

$$p_{ijk} = y_{ijk} - \bar{y}_{\cdot jk} \quad (1)$$

なる値の変動を考察した。 p_{ijk} は、 i 地点の風エネルギー特性を示す基礎的な数値である。本研究のすすめ方として、地形諸相から風を予測するモデルと全予測過程を組み立てる前に、個別にデータ各種の取り扱い法の検討を行う。本報では地形分類メッシュデータについて、説

明変数とするまでの取り扱いと、統計手法を呈示する。

2. 地形分類メッシュデータについて

国土数値情報は、国土整備事業の一環として、建設省国土地理院において作成された。国土数値情報の概要(1980)によれば、整備された情報はぎわめて多岐にわたる。これらのうち、本研究では下記3種類に限定して使用する。

加工が容易な連続量のデータとして、標高メッシュデータが第1のデータである。次に図形情報に近いものとして、地形分類メッシュデータの使用を試みる。第3に、土地利用(現況)メッシュデータを、必要に応じて使用する。土地利用状況が、風の強弱に影響されるとの想定により、風の結果といえる要因を逆用しようとするものである。これらのデータの単位は、いずれも3次メッシュ(約1 km×1 km)であるから、以後3次メッシュの数をもって、距離におきかえることがある。

本報では、地形分類メッシュデータの取り扱いについて述べる。

2.1. 地点の周辺地形範囲と地点別データの作成

* Estimation of small scale wind energy for agriculture by using mesh data 2. A method of multidimensional analysis of similarities using topographical mesh data.

** Shoko Hashiguchi, 農林水産省畜産試験場。

—1982年7月2日受領—

—1983年6月23日受理

第1表 地形分類データの主な分類基準

主 分 類	コード
大起伏山地	01
中起伏山地	02
小起伏山地	03
山麓地	04
大起伏火山地	05
中起伏火山地	06
小起伏火山地	07
火山麓地	08
大起伏丘陵地	09
小起伏丘陵地	10
火山性丘陵地	11
砂礫台地(上位)	16
砂礫台地(中位)	16
砂礫台地(下位)	16
ローム台地(上位)	14
岩石台地(上位)	17
岩石台地(中位)	17
岩石台地(下位)	17
扇伏地性低地	19
三角州性低地	21
自然堤防, 砂州	22
砂丘低地	23
湖 沼	24
河 川	25

注) 上のほか、県により、火山性扇伏地(12)、熔岩台地(18)など、やや特殊なものがある。

藤部・浅井の研究(1979)によると、(通常の)海陸風系は、ほぼ海岸線から20~40 kmの範囲にとどまる。また、吉野(1961)は、いくつかの地域について海風の侵入距離と方向を述べたが、概括すると侵入距離は10~40 kmとなる。更に吉野(前掲書)は、いくつかの局地的強風についての実態紹介のなかで、広戸風について、那岐山の南ろく10 km×10 kmの地域に吹く、と述べている。

ここでは風の種類には立ち入らないが、上の諸研究から、ある地点の風に影響するであろう周辺地形範囲として、半径10 kmでは過少と考える。したがって、該当地点を中心に、半径20 kmの範囲を目安として、データを用いることにする。

次に、集録するデータの単位地区が問題となる。地形

分類メッシュデータは、3次メッシュ単位で、地形の種類がコードで記されており、半径20 kmの地域範囲を示すデータ単位としてはマイクロにすぎる。しかし分類データの集約については一定の検討を必要とするので、データの収録は3次メッシュ単位のままとする。すなわち該当地点を含むメッシュを中心に、東西南北に20メッシュずつ方形をとる。したがって41 km×41 kmの範囲のものを、中央の該当地点がもつ地形分類データとして、以後地点別地形分類データと呼ぶ。第1表に分類内容とそのコードを示したが、方形内に海域を含むものがあるので、第1表に加え、海を示すコードを用いた。

2.2. 地形分類データ使用の意図

2.2.1. 予測モデルの考え方

本研究の最終目標はすでに述べた。簡単にいうと、メッシュデータから風の状態を予測することをめざす。予測をめざすとは、要因の解析が目的ではないという意味である。地形諸相が風に与える影響を定量的に調べようとするのであれば、土地利用データの使用は適切ではない。

更に、風と説明変数を結ぶモデルは、関数モデルではなく、予測モデルをめざす。この両者のちがいは、Draper・Smith(1966, 訳1967)の概念による。すなわち関数モデルとは「真の関数関係」であり、これを知ることが出来れば、もちろん予測は可能である。予測モデルとは「ある意味では非現実的ではあろうが、少なくとも今問題にしている応答と挙動の主要な特徴は再現してくれる線型予測模型」である。

上で引用した予測モデルの「非現実性」とは、風と地形との関係における既知の情報を無視する、という意味ではない。先に橋口ら(1980)は、イネいもち病の予測モデル作成にあたり、いもち病と気象要因との関係についての既往の知見が、モデル作成上重要なものであることを強調した。予測モデルについてのこの基本的な考え方は、ここでも同様である。

2.2.2. 地形パタンを内容とする説明変数

吉野(前掲書)によれば、谷が山地から平地にでた付近では、所によってはなほだしい強風が吹く。この例のように、風に影響する地形は、パタンとして記述されているものが多い。地形分類データ使用にあたり、たとえば大起伏山地が x 個連続する、小起伏山地と低地が交互にある、…、など主要なパタンを整理して分類することが可能であれば、データ数の問題はさておき、いくつかの明確な地形パタンを内容とする変数を使用することが

第2表 本研究で使した地形分類基準.

分 類	国土数値情報 のコード	統 合 した 新 コード
大起伏山地	01, 05, 09	1
中起伏山地	02, 06, 10	2
小起伏山地	03, 07, 11	3
山 麓 地	04, 08	4
合地, 丘陵など	12~18	5
低地, 砂地	19~23	6
湖沼, 河川	24, 25	6(7)
海 域		9

注) 湖沼, 河川は, 中央セルにおいてのみコード7とした.

できる. しかし, 風に異なる影響を与えるパターンは多岐にわたり, 事前に有効な説明変数を作成することは不可能に近いので, ここではその方式は採らない. しかし, 地形パターンを説明変数とする意義は大きいと考えるので, パターンを内容にもつ変数を, 統計手法により導出することを試みる. すなわち各地点のもつ地形分類データの組み合わせに対して数量を与え, その数量を変数の値とする. この場合, 数量の大小に対応してパターン内容がどのように変わるかは明確でないであろう. したがって変数内容の詳細な検討や結果の考察ができないという欠点があるが, 地形分類データを使用する意義は, 地形パターンを表す変数の使用にあると考える. 地点周辺の開放度, 起伏の程度などを表す説明変数も必要であるが, これらは内容の定義を明確にして, 標高データより作成する.

3. 地点別地形分類データの変換

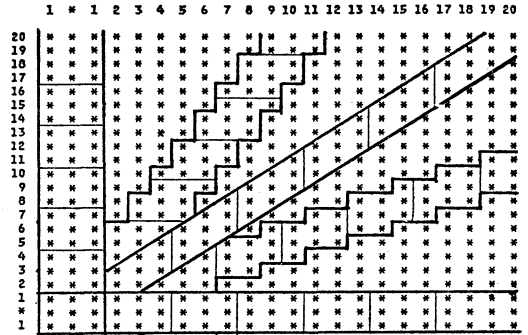
地点別地形分類データのデータ単位地区がマイクロにすぎことは, すでに述べた. 分類データの場合, 分類基準も当然問題となる.

3.1. 分類基準の統合

まず, 地表の粗度などは考慮しないこととし, たとえば第1表の大(中, 小)起伏山地と大起伏火山地は区別せずに扱うことにする. そこで, 第1表の分類を, 第2表にしたがい統合してコード変更を行い, これによって88個の地点別地形分類データのおきかえを行った. 以下, 地点別地形分類データとは, 第2表の分類基準によるものをさす.

3.2. 地点別地形分類データの縮約

3.2.1. 方形内データの取捨と8×13個のセル



第1図 地点ごとの地形分類メッシュデータ約1/4部分におけるデータの取捨とセルの作り方.

(1) 図中*印は, データの具体的数値に代わり記した, 2) 全体図においては, 上図左下端の方形が中央セルであり, これを中心として16方向に, 上と対称の形状で総数(16×6+1)個のセルが作られる.)

41×41の分類データの行列を, 以下のように縮約する.

1) 方形の中心を通して8方向に3列ずつ, 放射状にデータをとる. 第1図にその約1/4部分を示す. 第1図左下端の方形が, 全体図の中心に相当する.

2) 8方向の各々を, 13個のセルに分割する. 第1図にみられるように, セルのうちの多くは9個ずつのデータをふくむが, 12個, 11個あるいは6個から成るセルもある.

3) 8方向のうち, NE-SWの方向を1とし, 以後時計まわりで各方向を2, ..., 8とする. 更に各方向内で13個のセルに1, ..., 13の番号を付す.

3.2.2. セルを代表する地形分類コード

各セル内のデータを要約して, ひとつの分類コードで, そのセルの地形を代表させる.

各セルの座標を(a, b)とする. a=1, ..., 8; b=1, ..., 13である. 次の要約法による.

1) セル内9個(とは限らないが)のデータのうち, 最多数コードによって代表させる.

以下は, 同数のコードが2種類以上あるため, 1)で決め得ない場合の処理である. 隣接セルを代表するコードに依存するので, 各セルの処理順は, 中央セルから外側へ向かうことを原則とする.

2) 座標(a, 7)すなわち中央セルにおいては, セルの中心メッシュがもつコードを選ぶ.

3) (a, b) セルにおいては, $(a, b-1)$ または $(a, b+1)$ のセルを代表させたコードを除外する.

4) (a, b) セルのデータに $(a, b-1)$ または $(a, b+1)$ なる処理済みのセルのデータも加え, 最多数のコードを決める. ただし3) で除外したコードは復活させない.

5) セル内で, 各データの位置により重みをつけて1)にもどる. 中央位置のデータをもっとも重視(重み4)する. 全セル内各データについて, 1, 2および4の重みづけを定めたが, 殆どのセルでは4)までの方式で決めることができたので, 重みの呈示は省略する.

以上は, ある程度の数をもつコードには, 可能な限り, どこかのセルを代表させようとの考えを基本としたものである.

上の処理を終えると, 88地点のそれぞれが第1図の約4倍の各セルごとに, 分類コードを記した図をもつことになる. この図を $D_i, i=1, \dots, 88$, と仮称する. 図 D_i は, 各セルが所属する方向の番号1, \dots , 8を行番号として 8×13 の行列とすることができる. この行列を C_i と書く.

3.3. 地点間の類似性行列

地形分類データから地形パターンを表す変数を導出する, と先に述べた. D_i の作成までは, データ単位などの変更によるデータの縮約作業であるから, 地形分類データの使用目的に依存するのは, 以下の取り扱いであり, 二つの手法が考えられる.

いずれも各地点のもつ地形パターンに数量を与えるものであるが, ひとつは C_i をそのまま用い, 他は地点間の類似性から導く手法である. どちらも林・鮎戸(1976)および斎藤(1980)により多次元尺度法のなかに位置づけられている.

C_i の各行を個別に扱って, 内陸でのサンプル数を 49×8 とし, 各セルをアイテム, 分類項目をカテゴリーとして数量化Ⅲ類を適用, 地形パターンの数量化を行った. しかし, この数量を説明変数の値とするのは適切でないとの結論を述べるにとどめる. つまるところ, 上記のような扱いでサンプル数を多くしても, このデータから想定されるパターン数に比べ, サンプル数が少なすぎると思われる. したがって D_i を再縮約するか, たとえば中央部分9個のセルだけを用いる, などが必要であろう. すなわち数量化Ⅲ類の使用を前提とすれば, データの単位をより大きくするか, あるいは2.1.節で定めた周辺地形を実質上狭いものとしなければならない. このいずれをも避けるため, 以下のように地点間の類似性

による方法を用いる.

内陸立地点においては, 図 D_i の16方向のいずれからも風が同等の可能性で侵入すると考え, この考えに基づいて類似性行列を求める. しかし海岸立地点においては, 海風に重要をおく必要があるから, 類似性行列の作成方法は, 内陸と同じではない. そして, この行列の作成法のちがいは, その後の取り扱いに影響する可能性もある. 本稿では以後, 内陸立地の49地点を対象を限定して述べる.

3.3.1. 類似, 非類似と1, 0への変換

i 番目の地点がもつ行列 C_i の要素を

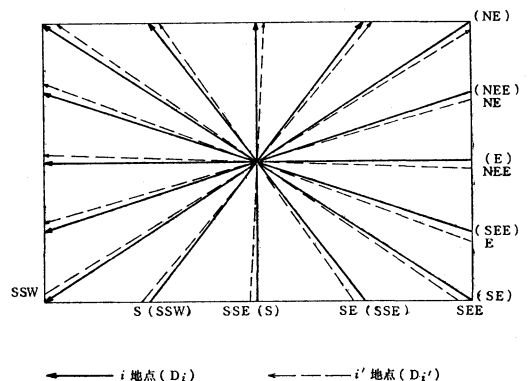
$C_i = (C_{iab}), a=1, \dots, 8, b=1, \dots, 13$, と書く. C_i と $C_{i'}(i, i'=1, \dots, 49)$ の, すべての a, b に対し

$$\begin{aligned} \delta_{ab}(ii') &= 1, C_{iab} = C_{i'ab} \text{ のとき} \\ &= 0, \text{ 上記以外 のとき} \end{aligned} \quad (2)$$

とすると, 特定の i, i' について0と1だけを要素とする 8×13 の行列が作られる. ところで C_i は, 図 D_i と対応するので, 式(2)は, D_i と $D_{i'}$ とを, 方向を合わせて重ね, 同じ位置にある二つのセルの地形分類コードが同じ(類似)なら1, 他を0としたものである.

3.3.2. $D_{i'}$ の回転による比較

風が吹いてくる可能性は, いずれの方向でも同じ, と考えたので, i と i' との互いに異なる方向の類似もみなければならない. すなわち D_i は固定したまま, $D_{i'}$ を $\pi/8$ (1方向)ずつ回転させ, 回転の度に $\delta_{ab}(ii')$ を作成すればよい. 第2図は $D_{i'}$ を1方向だけ回転させた場合であり, D_i の NEE-SWW, E-W, \dots , なる各方向に対し,



第2図 $r=1$ で D_i と比較するときの $D_{i'}$. (D_i の各方向を () 内に, 比較される $D_{i'}$ の方向をその下または左に, それぞれ記した. 矢印の方向はセル番号の昇順を示す.)

第3表 r の変化にともなう C_i と $C_{i'}$ の比較.

	C_i の行 番号 a	$C_{i'}$ の行 番号 a'	第 a' 行の要素の順 ¹⁾	
			$r \leq 7$	$r > 8$
$r=0, 8$	1	1	1 ~	13 ~
	2	2	1 ~	13 ~
	⋮	⋮	⋮	⋮
	8	8	1 ~	13 ~
$r=1, 9$	1	2	1 ~	13 ~
	2	3	1 ~	13 ~
	⋮	⋮	⋮	⋮
	8	1	13 ~	1 ~
$r=2, 10$	1	3	1 ~	13 ~
	2	4	1 ~	13 ~
	⋮	⋮	⋮	⋮
	8	2	13 ~	1 ~
$r=7, 15$	1	8	1 ~	13 ~
	2	1	13 ~	1 ~
	⋮	⋮	⋮	⋮
	8	7	13 ~	1 ~

1) i の第 a 行における要素 $(a, 1), \dots, (a, 13)$ と比較する i' の第 a' 行は $(a', 1), \dots, (a', 13)$ および $(a', 13), \dots, (a', 1)$ の2通りがある。このいずれであるかを示す。

$D_{i'}$ の NE-SW, NEE-SWW, ... がそれぞれ比較される。第2図図中の矢印の位置が13番目のセルに対応するので、 D_i における NE-SW 方向の各セル、 $b=1, \dots, 13$ に対し、 $D_{i'}$ では SSW-NNE のセルは $b=13, \dots, 1$ の順で同じ位置を占める。したがって、第2図に示されるような、1方向の回転による比較は、

$$\left. \begin{aligned} \delta_{ab}(ii'|r) &= 1, C_{iab} = C_{i'a+1b}, a=1, \dots, 7 \\ &= 1, C_{i8b} = C_{i'1h}, h=14-b \\ &= 0, \text{上記以外} \end{aligned} \right\} b=1, \dots, 13 \quad (3)$$

と書くことになる。

更に続けて、 $D_{i'}$ の回転と比較を繰り返すとすると、 $\delta_{ab}(ii'|r)$ は $D_{i'}$ の回転数 r に依存するから、 $\delta_{ab}(ii'|r)$ と書くことにすると、 $r=0, 1, \dots, 7$ のとき $q=8-r$ として

$$\left. \begin{aligned} \delta_{ab}(ii'|r) &= 1, C_{iab} = C_{i'a+rb}, a=1, \dots, q \\ &= 1, C_{iab} = C_{i'a-qb}, r \neq 0, \\ &= 0 \text{ 上記以外} \end{aligned} \right\} \begin{aligned} b &= 1, \dots, 13 \end{aligned} \quad (4)$$

が、式(2)と(3)をも含む一般式である。

$r=8, \dots, 15$ に対して、式(4)に準じて一般式を書くことができるが省略する。 $r=0, \dots, 15$ の各々に対し、 C_i と $C_{i'}$ とで比較される各行を第3表に示した。 D_i を固定するため、第3表の C_{iab} は、 r の如何にかかわらず、 $a=1, \dots, 8, b=1, \dots, 13$ なる順序とする。これらに対応し比較されるべき $C_{i'a'b'}$ の行 a' および行内の要素 b' の順序を記したものである。特定の i と i' に対し、 $\delta_{ab}(ii'|r)$ を要素とする 8×13 の行列、 $\Delta_r(ii')$ が、 r の変化にともない16個作られることになる。第4表に、イイヤマとハクバについて、 C_i と $C_{i'}$ および16個の $\Delta_r(ii')$ のうち、 $r=0, \dots, 3$ に対応するものだけを示した。

D_i における $(a, 7)$ セルは、 r の如何にかかわらず $D_{i'}$ の $(a', 7)$ セルと比較される。結局、すべての r と a に対し $\delta_{ar}(ii'|r)$ は同一比較を繰り返すことになる。第4表は、 D_i と $D_{i'}$ の中央セルが同じコードの地形(類似地形)であり、したがってすべての Δ_r の第7列の要素がすべて1となった例である。

3.3.3. 地点間の類似性行列

上述のように、49地点についての i と i' なる2地点ずつの組み合わせごとに、第4表のような $\Delta_r(ii')$ が16個作られる。この16個の行列を、ひとつの数値でおきかえなければならない。

特定の i, i' および r に対し

$$\begin{aligned} \delta_{ab}(ii'|r) &= \delta_{ab+1}(ii'|r) = 1 \\ (\text{または} &= \delta_{ab-1}(ii'|r) = 1) \end{aligned}$$

の場合、 D_i の a 方向と $D_{i'}$ の a' 方向において、連続する2組のセルの類似を意味する。おそらく、これらの方向における、2組の不連続なセルの類似に比べ、連続類似の方がより強く風の状態の類似を反映するであろう。そして、ふたつのセルでカバーする距離はほぼ6kmの程度であるから、ここで連続類似だけが風の状態に関わりをもつと仮定する。したがって、連続類似の数の総和をもって、 i と i' との各周辺地形の類似の指標とする。すなわち $\Delta_r(ii')$ の第 b 列と $b+1$ 列との内積を、 $d_{br}(ii')$ とすると

$$l_r(ii') = \sum_{b=1}^{13} d_{br}(ii') \quad (5)$$

第4表 $C_i, C_{i'}$ および $\Delta_r(ii')$ の例示.

	C_i (イイヤマ)	$C_{i'}$ (ハクバ)				
	3232116333343	2111116111111				
	3332216211342	3511116111111				
	1121156323331	4132226111111				
	1113116223432	3331236111111				
	1132256333441	6333336211111				
	2215556233536	2233236211111				
	2661266622223	3333336141111				
	3333356515322	1111116121111				
	$\Delta_r(ii')$					
	0000111000000	1000111000000	0011001000000	1010001000000		
	1000011011000	0011101011000	1110101011000	0110001111000		
	0100001000001	0001001000001	0000001000001	0000001000001		
	0000001000000	0001001100000	0001001100000	0001001000000		
$r=0, \dots, 3$	0010001000001	0010101000001	0010001000001	1100001000001		
	1100001100000	0000001000000	0010001000000	0010001000000		
	0000001000000	0001001010000	0001001000000	0001001000001		
	0000001000000	0000001010001	0000001010000	0000001000100		
	6	7	6	8	5 5	
					5 5	

注) $\Delta_r(ii')$ の下部の数字は $l_r(ii')$ と $m_r(ii')$ の値である. いずれも $r=4, \dots, 15$ については省略した.

は, $\Delta_r(ii')$ の行方向で1の連続する回数を示す. 更に中央セル近傍において, 中央セルをかこむような2組のセルの連続類似を

$$m_r(ii') = l_r(ii') + \sum_{a=1}^8 \sum_{b=5}^9 \delta_{ab}(ii'|r) \cdot \delta_{a+1b}(ii'|r) \quad (6)$$

の右辺第2項によって数えることができる. 上の第4表の各 $\Delta_r(ii')$ の下に, $l_r(ii')$ および $m_r(ii')$ の各数値を示してある.

i と i' における周辺地形の類似性を

$$s_{ii'} = \left\{ \sum_r m_r(ii') \right\} / 1984 \quad (7)$$

によって表す. ここに1984は, 式(7)の分子がとり得る最大値であり, すべての $\Delta_r(ii')$ の全要素が1であるとき得られる. したがって $i=i'$ のとき $s_{ii'}=1.0$ である.

$s_{ii'}$ を要素とする行列 $S=(s_{ii'})$ を, 類似性行列とよぶ. S は, 対角要素1.0をもつ 49×49 の対称行列である.

4. 類似性行列による各地点の数量化

4.1. 目的変数の種類について

吉野(前掲書)によれば, ある地点の風の状態に影響を及ぼす地形は卓越風向によって異なり, 同一の風向であっても風速によって影響する地形は違ってくる. したがってここでは, 地形諸相が風に与える影響に多面性, すなわち多次元解の存在を仮定する. このことは, 目的変数に対する示唆となる. 連続量1変数を目的変数

とすると, モデルのパラメーターに対し1組の解しか得ることはできない. 連続量多変数あるいは多群への各所属が目的変数のとき, 2次元以上の解を得ることができる.

ここでは目的変数として連続量多変数を前提とする. 具体的には式(1)の p_{ijk} (または \bar{p}_{ij} など) である. 本研究では, なお群所属という質的な目的変数の使用も考えられるので, この場合の地形パターンを表す説明変数の導出については別途考慮されねばならない.

4.2. 説明変数の導出

以下, 類似性行列 $S=(s_{ii'}), i, i'=1, \dots, L$ および式(1)の p_{ijk} の行列を Y とする. p_{ijk} は12変数であるが, ここでは p 変数とし, Y は $p \times L$ の行列と考える. ここでの目的は, S を手がかりに Y に対する q 個の説明変数の座標行列 X を得ることである.

特に Y を考慮することなく, 類似性を手がかりに $q \times L$ の行列を与える手法は, S を内積行列に限定しても, いくつか提起されている(斎藤, 前掲; 吉田, 1976). ここでも S を内積行列に対応させるが, 上記の手法によらず, Y をよく説明し得る変数の行列 X を求めようとするものである. 以下その手法を, 手順的に述べる.

1) 行列 S にもっとも近く, ランク $q (< L)$ をもつ行列を作り, これを $Z_1' Z_1$ と書く. S のもつ q 個の正の固有値を $\lambda_1 \geq \dots \geq \lambda_q \geq \dots \geq \lambda_L > 0$ とし, これらに対応する各固有ベクトルをとすると,

$$Z_1' Z_1 = \sum_{i=1}^q \lambda_i p_i p_i' \quad (8)$$

により得られる。S が非負固有値だけをもつとき——もし S が内積行列そのものであれば、当然非負固有値だけとなる-Rao(1964)により、ランク q をもち且つ S にもっとも近い行列が式 (8) で与えられることが知られている。そして、S が負の固有値をもつときは、式 (8) によって、その部分を除外することにもなる。ここに Z₁ は q × L 行列で、q 個の説明変数の測定値に相当するが、その存在を仮定するだけであり、実際にデータとして把握できるのは、式 (8) の Z₁'Z₁ である。

2) q 個の各説明変数について、それぞれが平均ゼロをもつように変換する。

現実には Z₁ を把握していないので、Z₁'Z₁ を変換の対象とする。すなわち

$$E = \left(I_L - \frac{ee'}{L} \right) \tag{9}$$

により

$$EZ_1'Z_1E = Z'Z \tag{10}$$

とすると、Z'Z は変換された内積行列となる。ただし I_L は L 次の単位行列、e' = (1, ..., 1) で L 要素のベクトルである。

3) 行列 Y に対し、X = H'Z の上の Y の全回帰項を最大にする q₁(≤ q) × L 行列 X を求める。

このような X を求めるための行列 H は、

$$|ZY'YZ' - \eta^2ZZ'| = 0 \tag{11}$$

を満たす固有ベクトルから成ることが、Rao (前掲) により示されている。奥野 (1974) は、この X を、回帰主成分と名づけた。

上の回帰主成分は、当然のことながら、Y と共に Z がデータとして把握されていることが前提である。しかしここでは、データとして具体的に把握されているのは Y および Z'Z だけであるから、これらにより X を求める方法を以下に示す。

式 (11) の解は、通常 ZZ' の q 個の固有値を λ_i、対応する固有ベクトルを q_i として

$$ZZ' = (\dots \lambda_i^{1/2} q_i \dots) \begin{pmatrix} \vdots \\ \lambda_i^{1/2} q_i' \\ \vdots \end{pmatrix} = Q^{-1} Q^{-1} \tag{12}$$

から

$$G = Q'ZY'YZ'Q \tag{13}$$

の固有ベクトルを経て求められる。したがってここでは、Z'Z の非ゼロ固有値と ZZ' のそれとが同じであること、Z'Z の固有ベクトルを p_i とすると

$$q_i = \lambda_i^{-1/2} Z p_i \tag{14}$$

なる関係がある (Rao, 前掲) ことを利用する。すなわち

$$Q' = \begin{pmatrix} \vdots \\ \lambda_i^{-1/2} q_i' \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ \lambda_i^{-1} p_i' Z' \\ \vdots \end{pmatrix} \tag{15}$$

であるから、式 (13) の行列 G は

$$G = \begin{pmatrix} \vdots \\ \lambda_i^{-1} p_i' \\ \vdots \end{pmatrix} Z'ZY'YZ'Z (\dots \lambda_i^{-1} p_i \dots) \tag{13'}$$

と、既知のもの積として書くことができる。

G は q × q の行列で、通常は p, q のうちの小さい方と等しいランクをもつはずである。G の固有値を大きい順に η₁² ≥ η₂² ≥ ... とし、これらに対応する固有ベクトルの行列を W とすると

$$\begin{aligned} X &= H'Z \\ &= W'Q'Z \\ &= W' \begin{pmatrix} \vdots \\ \lambda_i^{-1} p_i' \\ \vdots \end{pmatrix} Z'Z \end{aligned} \tag{16}$$

を得る。H' は得られないし Z も未知のままであるが、ここで必要とするのは X であり、式 (16) により導出

第5表 内陸20地点の1, 4月の y_{ijk} 値。

地点 番号	1月			4月		
	アサ (1~8)	ヒル (9~16)	ヨル (17~24)	アサ (1~8)	ヒル (9~16)	ヨル (17~24)
6	0.47	1.10	1.35	1.35	3.50	2.70
7	2.10	4.58	2.39	3.90	7.28	4.65
8	1.63	2.56	1.69	2.97	6.28	3.93
9	1.73	3.52	2.56	2.13	6.42	3.50
10	2.17	4.13	3.29	3.29	6.63	5.43
11	0.94	3.34	1.85	2.15	5.67	4.07
12	0.03	0.89	0.15	0.18	3.23	0.75
13	2.11	4.52	2.47	3.78	6.67	4.08
14	2.87	5.60	4.54	3.51	6.41	5.89
15	1.21	3.84	1.76	2.22	5.60	4.28
16	2.47	5.11	2.76	2.78	6.23	3.92
17	2.70	5.60	3.00	3.24	6.46	4.10
18	2.23	4.06	2.96	2.51	5.68	4.59
19	3.68	4.87	4.73	3.42	6.12	4.77
20	1.11	3.16	1.77	1.23	5.01	2.43
21	1.56	3.89	1.74	1.72	5.83	2.37
22	3.06	4.96	4.83	3.98	6.68	6.27
23	1.05	4.53	1.87	2.23	5.93	3.42
24	3.48	6.05	3.90	3.47	7.05	3.57
25	2.74	6.10	4.21	4.40	7.27	6.02
平均	1.97	4.12	2.69	2.72	6.00	4.04

第6表 $q=15, q_1=3$ による推定値 \hat{y}_{ijk} 値.

地点 番号	1月			4月		
	アサ (1~8)	ヒル (9~16)	ヨル (17~24)	アサ (1~8)	ヒル (9~16)	ヨル (17~24)
6	0.76	2.15	1.61	1.28	4.28	3.05
7	2.10	4.69	2.44	3.67	7.38	4.63
8	0.78	2.62	1.18	2.36	5.88	3.79
9	1.73	3.99	2.17	2.78	6.30	3.82
10	2.20	4.25	3.39	3.50	6.49	5.66
11	1.90	4.05	2.59	2.81	6.12	4.16
12	0.37	1.82	0.73	0.50	3.81	1.46
13	2.84	5.70	3.25	3.81	7.48	4.46
14	2.86	5.07	4.22	3.58	6.42	5.62
15	1.13	2.84	1.89	2.14	5.32	3.84
16	2.73	5.17	3.58	3.12	6.34	4.28
17	2.59	5.12	3.19	3.09	6.50	3.95
18	1.75	3.71	2.64	2.76	5.89	4.48
19	3.21	5.55	4.51	3.27	6.11	4.90
20	1.20	3.02	1.68	1.38	4.70	2.33
21	2.07	4.61	2.29	2.78	6.48	3.26
22	3.21	5.45	4.84	4.06	6.73	6.47
23	1.17	3.08	1.54	1.64	5.08	2.53
24	2.24	4.82	2.42	2.42	6.11	2.62
25	2.50	4.70	3.65	3.52	6.53	5.43

できる。そして式 (12) から

$$XX' = W'QZZ'QW = Iq. \quad (17)$$

すなわち X は q 個の互に無相関な成分から成る $q \times L$ の行列である。 η^2 の値の大きさの順にこのうちの q_1 個を選べばよい。この $q_1 \times L$ の行列を、区別して X_1 と書くことにする。

4) X_1 上の Y の回帰係数行列 B_1 を求める。

式 (17) から X_1X_1' も当然単位行列であるから

$$B_1 = YX_1' \quad (18)$$

である。そして Y の各値に対応する推定値の行列は

$$\hat{Y} = B_1X_1 \quad (19)$$

により求められる。

4.3. 数値例

上述のうちの主要な各式について数値例を示す。前報第1図に示した内陸49地点のうち、地点番号6~25の20地点を数値例の対象とする。

第5表に $y_{ijk} (= p_{ijk} + \bar{y}_{jk})$ を、比較のため第6表に式 (19) による \hat{y}_{ijk} を、いずれも1月と4月の値を例示した。第6表は、 $q=15, q_1=3$ による。計算の主要

第7表 類似性行列 S の部分行列.

地点 番号	6	7	8	9	10	11	12	13	14	15
6	1.00	0.05	0.07	0.09	0.05	0.08	0.07	0.08	0.04	0.07
7	0.05	1.00	0.05	0.10	0.05	0.10	0.12	0.05	0.04	0.09
8	0.07	0.05	1.00	0.10	0.04	0.08	0.06	0.05	0.05	0.09
9	0.09	0.10	0.10	1.00	0.07	0.22	0.08	0.10	0.06	0.20
10	0.05	0.05	0.04	0.07	1.00	0.08	0.07	0.06	0.04	0.07
11	0.08	0.10	0.08	0.22	0.08	1.00	0.10	0.09	0.08	0.16
12	0.07	0.12	0.06	0.08	0.07	0.10	1.00	0.04	0.03	0.09
13	0.08	0.05	0.05	0.10	0.06	0.09	0.04	1.00	0.06	0.06
14	0.04	0.04	0.05	0.06	0.04	0.08	0.03	0.06	1.00	0.05
15	0.07	0.09	0.09	0.20	0.07	0.16	0.09	0.06	0.05	1.00
固有値	2.50	1.27	1.06	1.04	1.02	0.99	0.96	0.95	0.94	0.93

注) S は 20×20 の行列であるがその $1/4$ を示す。固有値は S によるが、値の大きな順に10個を記した。最少値は0.65である。

第8表 各地点に与える3次元の数量.

($q=15$)

地点 番号	x_1	x_2	x_3
6	-0.2982	-0.2949	0.1839
7	0.1582	0.0494	-0.4772
8	-0.1203	-0.2906	-0.4202
9	-0.0207	0.0282	-0.2159
10	0.2141	-0.3022	-0.0585
11	0.0145	-0.0374	-0.0710
12	-0.5195	-0.0411	0.1740
13	0.2013	0.3190	-0.2326
14	0.2544	-0.1114	0.2221
15	-0.1248	-0.2714	-0.0637
16	0.1031	0.1967	0.1684
17	0.0685	0.2627	0.0388
18	0.0230	-0.1936	-0.0322
19	0.1868	0.1402	0.4598
20	-0.3130	0.0726	0.1687
21	-0.0413	0.2963	-0.1710
22	0.3868	-0.1983	0.2594
23	-0.2713	0.0592	0.0103
24	-0.1182	0.4684	0.0227
25	0.2163	-0.1517	0.0341
η_i^2	220.89	23.80	14.88
η_i^2 の比率 ¹⁾	0.6548	0.0706	0.0441

1) $Y'Y$ の固有和のうち η_i^2 の占める比率。

第9表 Y の各変数に対する回帰係数。
($q=15$)

		b_1	b_2	b_3
1月	アサ	3.2744	1.3673	0.9477
	ヒル	4.4001	2.5699	0.5381
	ヨル	4.4278	0.4416	2.0316
4月	アサ	4.0554	0.4150	-0.5916
	ヒル	3.5794	1.2227	-1.5896
	ヨル	5.2024	-1.7295	0.2938

な経過は、第7表以降で示した。

第7表に、類似性行列 S の一部分および S の非ゼロ固有値10個を記した。 S は20次の行列であり、非ゼロ固有値20個をもつが、全体として類似性は低く、したがって固有値の集中度も低い。 $q=15$ としたため、 S の固有値15個とそれらの固有ベクトルを用いて、式(8)により $Z_1'Z_1$ を、更に式(9)、(10)から $Z'Z$ を算出し、行列 G を経て式(16)の X に至る。

第8表は、式(16)による 15×20 の行列 X のうち、 $\eta_i^2, i=1, 2, 3$, に対応する3行について、それぞれの値と $\eta_i^2, Y'Y$ の固有和に占める η_i^2 の比率を示したものである。 η_i^2 の比率は q の値により異なるはずであるが、ここでは $q=15$ すなわちサンプル数20に対し、説明変数の数を比較的多くとったため、 η_1^2 の比率が高い。実際の適用の際には、 q の数の決定までに若干の試行錯誤を要するであろう。

第9表は、第8表における3次元の各成分と Y の12変数に対応する回帰係数行列 $B_1 = (b_1, b_2, b_3)$ の各ベクトルのうち1月と4月の要素の値である。先に述べた第6表の \hat{y}_{ijk} は、第5表との対応を容易にするため、第8、9表による $B_1 X_1$ の各値に \hat{y}_{ijk} (第5表の下段の値) を加算して示してある。

5. 適用にあたっての問題点

上記3.3.節以降においては、内陸立地点を対象とした。海岸立地点においても $D_r(ii')$ の作成は、内陸と同様に考えられよう。しかし D_i の回転を限定して、たとえば $r=0, \dots, 3$ にとどめる、などの変更が必要である

う。そしてこの程度の変更にとどまれば、4.章で述べた手法はそのまま適用可能であり、問題点の所在も内陸立地点と同等になる。

ここでは目的変数を y_{ijk} などの連続量とした。どのような目的変数を用いるにしても、標高データの利用による予測モデルの構築が重要なかなめとなる。そして、同一の目的変数に対し、地形分類データと標高データの両者を説明変数として用いる場合、類似性行列 S から導出される説明変数を、どこでどのように予測モデルに結びつけるかの問題が生ずる。

更に、未観測地点の値を実際に予測するにあたり、式(16)の X に相当する値をどのようにして求めるかが新たな問題となる。この問題の処理は、上でふれたような標高データの取り扱いなど予測の全過程の組み立て方により、異なるものとなる。

地形パターンを表す変数を導出し利用するにあたり、上のようになおいくつかの問題が残されているが、今後必要に応じて処理方法を提示する (GEP 82-I-3-1)。

文 献

- Draper, N.R. and H. Smith, 中村慶一 訳, 1966, 1967訳: 応用回帰分析, 森北出版, 234, 235.
服部文昭, 浅井富雄, 1979: 関東地方における局地風に関する研究, 第1部日変化を伴う風系の構造, 天気, 26, 595-604.
橋口渉子・加藤 肇・清水節夫, 1980: 農作物病害発生予測の一方法, 環境情報科学, 9-3, 56-63.
林知己夫, 飽戸 弘, 1976: 多次元尺度解析法, サイエンス社, 16.
建設省国土地理院, 1980: 国土数値情報の概要, 建設省国土地理院地理管理部.
奥野忠一, 1974: 回帰主成分とその数値例, 応用統計学, 4, 25-32.
Rao, C.R., 1964: The use and interpretation of principal component analysis in applied research, SANKHYA (A), 26, 329-358.
斎藤堯幸, 1980: 多次元尺度構成法, 朝倉書店, 7, 9-10.
吉田正昭, 1976: 心理統計学, 丸善, 240-258.
吉野正敏, 1961: 小気候, 地人書館, 145, 67, 64, 63.