原著論文

複数の順位情報に基づく評価関数の推定[†]

福田 亮治^{*1}・岡 薫^{*2}

評価関数があることが想定されるデータに対し、順位データが複数与えられた場合、そこに内在すると仮定する定量化された評価値を推定する手法を提案し、乱数を用いたシミュレーションと、グレー画像の明るさに関するアンケート調査をもとに、その有効性の検証を行った。この推定では、共通の評価値に一定の分布に従う観測誤差を加えたモデルに基づき、最尤推定により元の評価値を予測する手法を用いた。これにより、データにしかるべき多様性があれば、よい近似が可能であることが分かった。

キーワード：評価関数、順位データ、遺伝的アルゴリズム

1. はじめに

嗜好や使いやすさなどを解析したりするうえで、主観的な情報を定量化することがしばしば必要となる。一般にこれらの値を直接的に得ることは困難で、アンケートなどの不特定多数の対象から得られた情報から推測せざるを得ない。この場合、個人ごとのとらえ方の違いや調査方法の不統一性、被験者の集中力維持の困難などから信頼できる値を得ることは必ずしも容易ではない。この報告では、ある基準に基づき被験者が並べ替えた「並び順」がデータとして与えられたときに、そこに内在すると想定される評価関数を推定するためのモデルを提案し、その有効性に対して検証を行う。この手法は、[5]において手書き曲線の描画における分点の削減に必要な評価関数を構築するために用いたものである。

このような順位データは、個々の定量的な評価が記録されるものではないが、比較的容易に統一的な基準で評価を表現する手法であり、信頼できるデータであることが期待される。G.C.Reid[1]は、M.Friedman[2]やM.N.Brunden, N.R.Mohberg[3]の順位データに関する理論を用いて、経済的問題に関するある種の判断をする理論を構築している。この手法では、順序データを用いて χ^2 乗分布に従う統計量を構成し、その分布を用いた統計的判断を行っている。そのほか、Wilcomの並べ替え検定[4]などのノンパラメトリック

な統計手法は順位に関わる事象を並び順の全体のなかでとらえるものであり、順位データは古くから注目されている情報の形態であるといえる。

X を有限集合とし、 $f(x)$ を X 上で定義される関数とする。すべての順位はこの評価関数をもとに行われるが、 $x \in X$ ごとの観測には独立な誤差が加わり、そのために個々の順位が必ずしも同じでないとする。このとき、仮に2つの要素 $x, y \in X$ に対して特定の一人が $f(x) < f(y)$ であると判断した場合、これが明らかな違いなのか微妙な違いなのかは、1人分のデータからは知るすべがない。しかし、複数の人がこの並べ替えを行ない、ほとんどすべての人が同じ順番に並べた場合には、 $f(x)$ と $f(y)$ には十分に大きな差があることが予想され、逆にほぼ半々の人が逆の順番に並べた場合は2つの値は近いものと考えることができる。この考え方を元に評価関数を構成するためのモデルを次のように設定する。

1. $X = \{x_1, x_2, \dots, x_n\}$ を有限集合とする。
2. $f(x) : X \rightarrow \mathbf{R}$ を共通の評価関数とする。
3. $\{g_j\}_{j \leq n}$ を標準正規分布に従う独立な確率変数列とし $G_j = f(x_j) + g_j$ を $f(x)$ の観測を想定した確率変数とする。

個々の並べ替えデータは、 $\{G_j\}_{j \leq n}$ の実現値に基づく並べ替えであると考える。仮に $x, y \in X$ に対して $f(x) = f(y)$ であれば、 $f(x) < f(y)$ であると判断する確率は1/2である。逆に、 $f(x) - f(y)$ が十分大きい場合 $f(x) < f(y)$ と判断される確率はほぼ0に近い(差が4.25であれば確率は0.001程度)。そこで、実際に得られた複数の順位をもとに基準となる順位をつくり、要素の組み合わせ (x_j, x_k) に対して、順位データと基準

[†] Estimation of Evaluation Function using Data of Several Sort Results

Ryoji FUKUDA and Kaoru OKA

^{*1} 大分大学工学部

Faculty of Engineering Oita University

^{*2} 九州工業大学 情報工学部

Faculty of Computer Science and Systems Engineering Kyushu Institute of Technology

となる順位とが異なる個数 $r_{j,k}$ を調べる．並べ替え時の順位の違いの頻度は $f(x)$ の値により定まるので，上記モデルに基づく尤度の和が最大となるように，この関数の推定を行う．この場合，正規分布は原点を中心とした左右対称の分布であるので，初めに与えた基準となる順位に誤りがあっても，この推定に対する影響は少ないと考えられる．

実際の推定には2つの要素の組み合わせを複数考え，対応する順位の逆転個数に対応する対数尤度の和を用いている．しかしこれらの事象は一般に独立ではなく，対数尤度の和が全体の対数尤度にはならない．そのほか，いくつかの場面で近似的な手法を用いている．そこで4節では，これらの近似的，代替的处理の影響なくモデルに対する推定が有効に機能するかを検証するために，モデルに従ったデータを乱数を用いて作り，この手法により関数 $f(x)$ の推定をし，その精度の検証を行う．

さらに5節では，数枚の単色のグレー画像を明るさの順に並べ替えたデータを用いて，人の感じる画像の明るさの推定を行う．明るさの近い画像に対して人間が常に正しい判断をすることは困難であり，このモデルで想定している状況と近い例になっていると考えられる．さらにグレー画像の明るさは，Lab表色系のLの値と比較的相関が強いと考えることができるので，この値を $f(x)$ と考えることで，関数値が客観的に得られる状況である．人の手による並べ替えには，モデルで想定している誤差のほかに，操作の間違いや残りなどさまざまな要素が付加される可能性があり，それらを吸収して有効な解析ができるかどうかを確かめることを目的としている．

2. 推定に用いるモデル

2.1 Error Sort Model

$X = \{x_1, x_2, \dots, x_n\}$ を有限集合とし， X 上に関数 $f(x)$ が定義されているとする．この関数が順位データをもとにした推定の対象となる関数であり， X の各要素に対して並べ替えの基準となる評価を与えるものとする．被験者はこの関数に誤差が加わった次の確率変数列を観測し，それに基づいた並べ替えを行っているものとする．

$$\begin{aligned} \{\mathbf{G}^{(i)}\}_{i \leq N} &= \{\{G_j^{(i)}\}_{j \leq n}\}_{i \leq N} \\ &= \{\{f(x_j) + g_{i,j}\}_{j \leq n}\}_{i \leq N} \end{aligned}$$

ここに， $\{g_{i,j}\}_{i,j}$ は標準正規分布に従う独立な確率変数列とする．

各 $i \leq N$ ごとに $\{G_{s_{i,j}}^{(i)}\}_{j \leq n}$ が小さい順に並ぶように，番号列

$$\mathbf{S} = \{\{s_{i,j}\}_{j \leq n}\}_{i \leq N} \quad (1)$$

を定める．与えられるデータは，被験者の基準により X の要素(番号)を並べ替えた整数列であるが，上のように共通の基準に誤差が加わったものが根拠となっていると仮定する．このように誤差を含んだ評価関数をもとに並べ替えられたデータのモデルを **Error Sort Model** と呼ぶ．

2.2 2要素間の順位と評価関数の差

与えられた並び順 $\mathbf{S}$ の逆変換に当たる整数列を

$$\{\{s_{i,k}^* = j \text{ such that } s_{i,j} = k\}_{k \leq n}\}_{i \leq N} \quad (2)$$

と置く．(x_j が並べ替えられた列の何番目に来ているかを表す．)ここで， X の要素の並び順については，標準的なものが与えられていて，被験者の並べ替えはこの順位から大きくはずれないと仮定する．説明を単純にするために(必要なら与えられた番号付けを変更することにより)この節では X の要素はあらかじめ標準的な順序に並んでいるものとする．

X の異なる要素 $x_j, x_k (j < k)$ に対して

$$r_{j,k} = \#\{i : s_{i,j}^* > s_{i,k}^*\}$$

により標準の並び順とは異なる順位になっているデータ数を表す．($\#\{\dots\}$ は集合 $\{\dots\}$ の個数を表す．)

$$s_{i,j}^* > s_{i,k}^* \Leftrightarrow f(x_k) - f(x_j) < (g_j - g_k)$$

であり， $g_j - g_k \sim N(0, 2)$ であるので，全体数を N とする場合に相互の独立性を仮定すれば

$$Pr(r_{j,k} = m) = {}_N C_m p^m (1-p)^{N-m},$$

$$p = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{f(x_k) - f(x_j)} e^{-\frac{x^2}{4}} dx$$

により確率値が定まる．

一般に全体数 N の2項分布に対して，実現値 r_0 を得た場合 (N 回の独立試行中，対象の事象が r_0 回起こった場合) 確率 p の最尤推定値は $\frac{r_0}{N}$ である．観測された2つの関数値に付随する誤差に対応するエラー関数を

$$Er(x) = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^x e^{-\frac{x^2}{4}} dx$$

と定めると，

$$f(x_k) - f(x_j) = Er^{-1}(1 - \frac{r_0}{N})$$

により局所的に最適な関数値の差が得られる．

$\varepsilon > 0$ を0に近い正の数とする． α を，標準正規分

布に従う確率変数 X に対して $Pr(X > \alpha) = \varepsilon/2$ を満たす実数とする。正規近似による 2 項分布の確率値 p の推定区間は、母分散 σ^2 が既知と考えると

$$\left[\frac{r_0}{N} - \frac{\alpha\sigma}{\sqrt{N}}, \frac{r_0}{N} + \frac{\alpha\sigma}{\sqrt{N}} \right]$$

で与えられる。 σ^2 の理論値は確率値 p を用いて $p(1-p)$ とあらわされるので

$$\sigma \leq \sqrt{\max_{p \in [0,1]} p(1-p)} = \frac{1}{2}$$

が成り立つ。したがって信頼区間の幅は p の値によらず $\frac{\alpha}{\sqrt{N}}$ を超えないことが分かる。与える並び順は標準のものから大きくはずれないことを想定しているので、 r_0 は比較的小さい数であることが予想される。そこで $\frac{r_0}{N} - \frac{\alpha\sigma}{\sqrt{N}} > 0 \iff r_0 > \sqrt{N}\alpha$ であることを、信頼できる解析のための一つの基準と考えて、 ε ごとに決まる $\lfloor \sqrt{N}\alpha \rfloor + 1$ を安定逆順数と呼ぶことにする。ただし $\lfloor \sqrt{N}\alpha \rfloor + 1$ が $N/2$ を超えるときは、安定逆順数は定義できないものとする。 ε と N の値を変化させたときの安定逆順数を表 1 に示す。

理論的には(大数の強法則を用いて)、 $p > 0$ であれば、十分大きな N に対して、逆順の数が安定逆順数を超える確率が 1 であることが分かるが、関数値の差が大きいときは現実的な回数では実現しない。

この報告で行っている推定では、一組の要素の対の逆順の個数を用いて関数の差を推定し、それらにより全体の関数値を構築する方法を用いている。したがって、逆順が 1 つもないときにはこの手法による推定はできない。また、上で述べたように逆順の個数によっても処理の精度が変わることが予想される。そこで、逆順の個数に関して次のような概念を導入する。

ある要素間の逆順の個数が r_0 ($r_0 > 0$) を上回る場合にその 2 つの要素が接続していると考え、 X 要素間に同値関係を定めることができる。この同値関係に関して全体が一つの同値類になることを、「 r_0 -連結」と定義する。 r_0 -連結になる r_0 の最大値を最大連

結逆順数と呼ぶ。最大連結逆順数が比較的大きな値であるとき、よい推定ができることが期待できる。逆順の個数は、個々の並べ替えの違いによるもので、並べ替えの順序は程よく食い違っていたほうがよいことになる。

3. モデルに基づく推定手法

3.1 標準順位と逆順の個数

2 節では説明の簡略化のために X の要素が標準の順位と同じ順番で並んでいるとして、議論を進めてきたが、実際の処理ではデータに基づいて以下のような方法で標準の順位を決定する。

1, 2, ..., n を並べ替えたデータ(2.1節(1))に対して、2.2節(2)で定めた $\{s_{i,j}^* \mid j \leq n\}_{i \leq N}$ を用いて、順位の平均を次で定義する。

$$a_k = \frac{1}{N} \sum_{i=1}^N s_{i,k}^*$$

この値をもとに $\{a_{s_j}\}$ が単調増加になるように添え字指数を並べ替えたものを $\{s_j\}$ と置き、これを標準の順位とする。このとき ($j < k$) に対する逆順の個数は次のように与えられる。

$$r_{j,k} = \#\{i : s_{i,s_j}^* > s_{i,s_k}^*\}$$

3.2 評価関数の初期推定

簡単のために、 $x_{s_k}, x_{s_{k+1}}$ ($k=1, 2, \dots, n-1$) がすべて k -連結 ($k \geq 1$) であるとする。このとき、各 $j=1, 2, \dots, n-1$ に対して

$$d_j = Er^{-1} \left(1 - \frac{r_{j,j+1}}{N} \right)$$

を定めることができ、この値を隣接 2 点間の関数値の差と考える。端点での値を 0 として

$$f(x_{s_1}) = 0, \quad f(x_{s_k}) = \sum_{j=1}^{k-1} d_j \quad (3)$$

によりすべての関数値を決定する。

標準の順位は順位の平均をもとに作られているので、隣り合うものの逆順の個数が一般的に高くなりやすく、 k -連結であれば隣り合うものがすべて連結であることが多い。しかし、一般には全体が連結であっても隣り合うものが連結でない場合もあり得る。連結性がまったく無い場合を含めて、隣り合うものに連結でないものがあることを想定した次の手順で初期の関数値を決める。この手順において、なるべく信頼のおける組み合わせを用いて初期の関数を決めることを目指している。

表 1 安定逆順数

$\varepsilon \setminus N$	10	20	30	50	100	200	500
0.5	3	4	4	5	7	10	16
0.2	5	6	8	10	13	19	29
0.1	-	8	10	12	17	24	37
0.02	-	-	13	17	24	33	53
0.01	-	-	15	19	26	37	58

0. 最大連結逆順数が1以上であれば、 m を最大逆順数とし、そうでなければ $m=1$ とする。
1. $f(x_{s_1})=0$ とする。
2. 標準の順位で1番目の要素を始点として $(x_{s_1}, x_{s_2}), (x_{s_1}, x_{s_3}), \dots$ の順で接続を調べ(対応する逆順数が m 以上であれば接続すると考える)、初めに接続した要素 x_{s_r} に対して(3)の方法で関数値 $f(x_{s_r})$ を定める。
3. 同様に関数値 $f(x_{s_{r'}})$ 、($r < r'$)を定め、後に接続する要素が見つかる間繰り返す。
4. すべての関数値が定義されていれば終了する。
5. 関数値が未決定の要素の中で、標準の順位が一番早いものを固定する。固定された要素とすでに関数値の決まっている要素の接続を調べ、接続があれば、(3)の方法で関数値を決める。(未決定の要素すべてを対象に繰り返す。)
6. 4. に戻り十分な回数(要素数と同じ回数)繰り返す。
7. 終了しない場合は、(この場合連結ではない)(3)の方法を標準の順序で隣り合う要素に対して初めからやり直し、連結していないところは固定の十分大きな差(実際には5.0を使用)を与えすべての関数値を決める。

3.3 推定の改善

3.2節の推定において、仮の並び順での隣接する要素間の逆順の個数のみを使って関数の近似を行った。一般に逆順になる組み合わせはほかにもあり、0番目の手順で定めた m に対して、 m -連結となるすべての要素の組み合わせに対して、その逆順数を考慮した推定を行う。これら全ての関数値の差を最尤推定すると、一般に同時に満足する関数値を設定することはできない。そこで、これらを対象とした尤度を考えた処理をすることが有効と考えられるが、2つの要素の組み合わせに関わる事象は、一般に独立性がないので尤度を精密に表現するとかなり煩雑な処理が必要になる。そこで、(2つの)要素の組み合わせごとに尤度を計算し、その対数の和を用いて推定を行う。具体的には、遺伝的アルゴリズムに準じた以下の方法で推定を行う。データの並び順が比較的安定している場合は、改善手法を用いても初期推定から大きく変化しないことが予想されるが、一般論として初期推定の手法では情報のすべてを用いていず、付加された情報によりどの程度推定が改善されるかを調べる必要がある。

1. 3.2節の方法で初期の関数値を定める。
2. 1. で定めた関数値の各値に平均0標準偏差

0.1の正規乱数加え、微小量変化させた関数を M 個作る。 M を遺伝子数、それぞれの関数を遺伝子と呼ぶ。

3. 対象となる(逆順数 m 以上の)すべての要素の組み合わせ $x_j, x_k \in X$ に対して、モデルの誤差分布に基づき逆順になる確率値を求め、与えられた逆順数に対応する2項分布の対数尤度の和を計算する。
4. すべての遺伝子に対して、3. の対数尤度の和を計算し、大きい順に並べる。
5. 上位3個の遺伝子は変更を加えないものを残す。残りの遺伝子を均等に4つのグループに分け、最上位の要素から対応する個数をコピーする。
6. 1番目のグループには、3次元部分空間に制限したニュートン法を用いた改善を行う。
7. 2番目のグループには、勾配法による改善を行う。
8. 3番目のグループには、遺伝子の各成分に小さい標準偏差(0.05)の正規乱数を加える。
9. 4番目のグループには、2. とおなじ初期化をする。
10. 4. に戻り必要回数繰り返す。

具体的な処理においては処理の繰り返し回数を10回と遺伝的アルゴリズムとしては少ない回数で打ち切っている。この推定の主要部分は6番目、7番目の解析的な推定であり、最小値以外の極小値がなければ、これらの手法だけで(ニュートン法や勾配法としては1ステップ30回程度の繰り返しがある)十分推定されると考えられる。極小値が複数ある状況を想定してこれを回避するためこれ以外の処理も加えている。実際の推定では図1(対数尤度 $\times(-1)$ の推移)に見るように10回程度でかなりの改善がされていると期待できる。この報告の主旨は手法の評価であり、より正確な推定をすることではないので、計算時間節約のため、この推定における繰り返しを10回に固定し解析を行う。

図1 尤度 $\times(-1)$ の推移(例)

4. 乱数を用いたシミュレーション

前節で述べた手順には、独立性のない複数の事象の対数尤度の和をとるなど、処理を単純にするために近似的な方法を用いている部分がある。また、入れ替わり回数 n の比は大数の法則によりいずれその確率値に収束するが、個数が少ないときには必ずしも良い近似をしているとは言い難い。そのような状況下で、この手法による推定がどの程度本来の関数を再現できるかを、乱数を用いたシミュレーションで調べる。乱数の生成には、無理数の演算における少数部の挙動を用いた疑似乱数(H.Sugita[3])を用いている。

4.1 データ作成

推定の目標となる関数の値 $\{y_j\}_{j \leq n}$ と、それを観測したものを想定した値 $\{x_j^{(i)}\}_{j \leq n, i \leq N}$ を、正規乱数を用いて次のように作成する。

1. 標準正規分布に従う乱数を n 個発生させ $\{y_j\}_{j \leq n}$ とする。
2. $\{g_j^{(i)}\}_{j \leq n, i \leq N}$ を $N(0, \sigma^2)$ に従う正規乱数(確率変数としては独立同分布)とし、 $x_j^{(i)} = y_j + g_j^{(i)}$ により $x_j^{(i)}$ ($i \leq N, j \leq n$)を定める。

各 $j \leq N$ に対して、上の手順で得られた数列の番号を $\{x_j^{(i)}\}_{j \leq n}$ の大きさの順に並べ替えたものを $\{s_j^{(i)}\}_{j \leq n}$ とする。このデータのみを用いて、前節で述べた方法で $\{y_j\}_{j \leq n}$ の推定をする。

表2 最大連結逆順数

N	標準偏差	最大連結逆順数 (平均)	安定逆順数 ($\varepsilon = 0.2/0.5$)
50	0.1	0.01	10 5
	0.2	0.56	
	0.3	2.2	
	0.5	6.45	
100	0.1	0.02	13 7
	0.2	1.24	
	0.3	4.6	
	0.5	13.43	
200	0.1	0.05	19 10
	0.2	2.37	
	0.3	9.18	
	0.5	25.2	

4.2 最大連結逆順数

前節の方法でデータを作成した場合、誤差の標準偏差である σ の大きさにより、推定能力を左右する最大連結逆順数が影響を受けると考えられる。表2はデータ数と誤差の標準偏差ごとの最大連結逆順数の平均(サンプル数1000)である。この表にあるくらいのあまり大きくないデータ数では、高い精度での安定性を得ることは難しく、特に標準偏差が0.1の場合は1-連結性も得られないことが多々あることが分かる。

4.3 推定誤差

推定対象作成時の標準偏差1.0に対して、連結性があまり期待できない標準偏差0.1の場合と、比較的連結になることが多い標準偏差0.3の場合とで、推定時の誤差の比較を行う。図2、3は、データ数50、推定対象作成時の標準偏差1.0誤差の標準偏差0.1、0.3で作成したデータを用いて、遺伝子数50、10世代の推定を行った一例である。

双方とも推定対象、順位の平均、初期推定、詳細推定に対して、最大値と最小値を一次変換でそろえたものをグラフ化している。この推定の目的は判断の基準となる評価関数を推定するものであるが、この手法で

図2 シミュレーションの推定例 $\sigma=0.1$

図3 シミュレーションの推定例 $\sigma=0.3$

は絶対的な大きさを推定することはできない。そこで、相対的な大きさの比較をしやすようにこの形で表現している。この場合全体的なグラフの形状が同様であれば、正しい推定ができていると考えられる。特に中央部に誤差が積算される形となるので、グラフにおいては中央部付近のグラフの離れ方が全体としての推定の精度を、場所ごとの傾きの変化が同じかどうかが局所的な推定の精度をそれぞれ表していると考えられる。

これら2つのグラフには、共に初期推定と改良後の推定に大きな差は見られない。これは、逆順数が小さいため推定に使われている逆順の情報に隣接しているものが中心になっていることが原因と思われる。誤差の標準偏差が0.1のデータでは1-連結でなく、0.3のデータは1-連結(最大逆順連結数は1)になっている。0.3の推定では、順位の平均に比べ明らかによい近似になっている。0.1の推定においても、全体的に近づいていることに加えて、値の変化の形状が本手法のほうがよく似ていることが確認できる。全体が連結であれば、グラフの形状をほぼ忠実に再現する推定が行われ改良推定の必要がないこと、連結でない場合のグラフの形状のずれは改良推定では回復を期待できないことが分かる。図4は、それぞれ100回ずつのシミュレーションにおける推定に対して、 $\sigma=0.1, 0.2, 0.3$ における100回のシミュレーション結果を用いた推定対象との偏差の99パーセントの信頼区間である。比較対象である平均順位の偏差に対する信頼区間は、幅が広いだけでなく、信頼区間に0を含まないことがある。逆順数が十分であり全体が連結である場合には、よい近似ができていると考えられる。

図4 偏差の99%信頼区間

5. グレー画像の明るさの推定

人が主観的に判断するもので、客観的な数値的評価が既に与えられているものとして、色の明るさが挙げられる。この解析ではLab表色系のLの値を色の明るさを定量化した値として比較を行う。色の種類の違いによる印象の変化などを考慮して、調査にはグレーの単色画像を用いる。複数の画像を明るさの順に並べ替えたデータをもとに、提案手法に基づき明るさを推定しLの値と比較する。

5.1 画像の作り方

単色の色の明るさの違いは、画像を画面上で並べて比較するとかなり正確に判断をすることができる。予備実験としてRGBの3色の値がすべて等しいグレー画像256種類から一様乱数を用いてランダムに15個選んだものをもとに、画像を張り付けることのできる文書編集ソフト上で、実際に色を並べて表示することで並べ替えを行ったところ20名以上の被験者がすべて、正解と同じ並び順に並べ替えた。提案手法では、「標準の並び順と逆になるものが存在する」という関係で連結になっている状態が望ましいが、この場合はほとんどすべての点が孤立点になっている状態で、値の差を推定することができない。

そこで、RGB値(3色の値はすべて同じ)の取りうる範囲0~255をほぼ1/3の85~169に制限し、その中から一様乱数を用いて選んだ15色を用いた。選ばれた色の80×120(ピクセル)の矩形を、3×5並べたBMP画像を作り調査に用いた(図5)。選ばれた15色には、選んだ順にA~Oの記号を対応させ、矩形の下の部分に表示させた。さらに矩形の境界には太さ2ピクセルの青い線を引き、隣り合う色の変化がすこしわかりにくくなっている。

図5 調査に用いた画像

5.2 調査(データ収集)の方法

順位情報の取得は、15色の色が配置された画像を見ながら、テキストファイルに記号を書き込む形で行った。提示する画像は一つのBMP画像であり、画像を変更することなく答えてもらったので、その並び順により比較しやすさが変わることが予想される。そこで、画像の並び順は乱数により決め、基本的にすべての人が異なる並び順の画像を用いて並べ替えを行った。選ばれた色には全く同じ色はなかったが、色の判断がつかない場合は表示毎で上のものの優先、同じ段であれば左のものを優先してもらうことにし、つけているアルファベットの順による影響を排除した。また、調査に先立って目的を十分説明し、能力を測ることが目的ではないことを確認したうえで、基本的に直観に基づいて答えてもらった。

調査対象は10代後半から20代前半を中心とした30名で、高校生と理系の大学生、大学院生(工学部、情報工学部)がその大半である。

5.3 解析結果

データ数30はすべて矛盾(欠落や重複)のないものであり、得られた順位情報をもとにして得られる最大連結逆順数は1であった。安定逆順数($\epsilon=0.5$ で4)には及ばないものの、連結にはなっているので、全体をデータに基づいて推定することができる。(非連結の部分は一定の関数値の差が与えられ、推定時にその差が変更されることはない。)この情報を用いて、遺伝子数50世代数10で推定を行った結果が、図6である。シミュレーションの時と同様に、推定対象にあたる明るさ(Lab表色系のL)と順位の平均、初期推定、改良推定の値を、最大値と最小値がLの値と同じになるように変換したものを、標準の順序(順位の平均)に関して暗い順)にならべたものを表示している。したがって、グラフの両端はすべての値が一致し、違いは

図6 明るさの推定結果

中央部のみに見られる。この推定では、順位の平均を含めて全部がLの値を比較的忠実に近似している。順位の平均も比較的よい推定になっているので、推定の大幅な改善にはなっていないが、初期推定、改良推定は共にLの値のグラフの形状をほぼ忠実に再現している。

グラフの中心部を拡大してみると、順位の平均のものに比べて推定が改善されていることが分かる。変化としては微妙なものであるが、明るさとLの値の関係についての細かい議論なしに本手法の有効性を主張できるものではないが、グラフの傾きの変化を含めて本手法の改良推定が、一番Lの値に近いものになっている。

6. まとめと今後の問題

本報告では、ある基準に従って複数の人が並べ替えたデータをもとにその基準に対応する評価関数を推定する手法を提案し、モデルに基づくランダムなデータと、グレイ画像数枚を明るさの順に並べたデータを用いて、その性能を調査した。この手法が有効に働くためには、標準となる順番とは異なる順番で並べられたもの(逆順数)がある程度必要であり、逆順を用いて定義した連結性のあるなしで、その推定精度に差がみられた。また、乱数を用いたシミュレーションでは、本手法による推定は順位の平均から求めた推定よりも対象となる評価関数をより忠実に再現することが分かり、画像の明るさに関する調査では、LabのLの値のグラフの形状をほぼ忠実に再現することができた。

本論文では、標準の順位で隣り合うものだけの情報を用いた初期推定と、有効な逆順数を可能な限り用いた改良推定を提案しその性能を調べたが、改良推定では初期推定からの大きな改善は見られなかった。逆順数に関する連結性のあるところでは、初期推定が十分に良く、改良推定なしで十分な精度があると考えられるが、連結性の無いときの推定の悪さを改善しようものでは無いと考えられる。これは、逆順の比較的起こりにくい状況での調査だったことが原因と考えられるが、逆順がかなり多い場合では有効である可能性を残す。例えば、今回考察の対象としなかった不完全データ(対象となる要素のなかの上位の一部だけを並べたもの)においては、設定によって順位に大きな変動が予想され、改良推定の活用を含め、本手法の適用の可否を調べていきたいと考えている。

また、連結性が期待できないような状況では、現在想定している情報からでは推定に限界があると考えられる。例えば「順位を付ける場合にどの程度判断に迷うか」という情報など、付加的な情報を加えることで改善が期待できるが、この場合データ作成の簡易性を損なう恐れが大きい。個々のデータに、被験者への負

担なく「どの程度迷ったか」がわかる情報を付加することができれば、さらに多くの状況下でこのモデルを適用することができると思われる。

謝辞

この論文を作成するにあたってお世話になった、九州工業大学大学院情報工学研究院の岡崎悦明教授、本田あおい准教授に深く感謝の意を表します。

参考文献

- [1] C.R.Cavin, Analysing Rankings, with an Application to the Financing of Small Entrepreneurial Firms, The Economic Journal, 100(1990), 200-205
- [2] M. Friedman, The Use of Ranks to Avoid the Assumptions of Normality Implicit in the Analysis of Variance, Journal of the American Statistical Association, Vol. 32(1937), 675-699
- [3] M.N. Brunden and N.R. Mohberg, The Benard - Van Elteren statistic and nonparametric computation, Communications in Statistics, Simulation and Computation,

Vol B 5(4) (1976), 155-162

- [4] E.L.Lehmann / 原著 鍋谷清治・刈屋武昭・三浦良造 / 共訳, ノンパラメトリックス (順位に基づく統計的手法), 森北出版
- [5] K. Oka and R. Fukuda, Reducing Points in a Hand-written Curve (Improvement in a Note-taking Tool), Proceedings of 15th Asian Technology Conference in Mathematics.
- [6] H.Sugita, Psuedo-random number generator by means of irrational rotation, Monte Carlo Methods and Applications, Vol.1-1, VSP(1995), 35-57.

(2011年2月2日 受付)

(2011年6月8日 採録)

[問い合わせ先]

〒870-1192 大分市大字旦野原700

大分大学工学部福祉環境工学科

福田 亮治

TEL : 097-554-7860 7860(内線)

FAX : 097-554-7860

E-mail : rfukuda@oita-u.ac.jp

著者紹介

ふくだ りょうじ
福田 亮治 [正会員]

大分大学工学部福祉環境工学科准教授、博士(数理学)、視覚障害者支援を中心とする福祉情報工学および非加法的測度を用いた主観的評価問題に関する研究に従事。日本知能情報フェジ学会、電子情報通信学会会員。

おか かおる
岡 薫 [非会員]

九州工業大学情報工学部システム創成情報工学科卒業、同大学院情報科学専攻在学(修士2年)、非加法的測度を用いた主観的評価問題に関する研究に従事。

Estimation of Evaluation Function using Data of Several Sort Results

by

Ryoji FUKUDA and Kaoru OKA

Abstract :

We propose a novel quantization method for human likelihood or preference using several sorting results. We estimate an evaluation function by assuming that the data is sorted on the basis of observed function values and that the sorting order varies from person to person according to the observation error for the function. We present tests to evaluate the proposed method by using random data and sorting data for gray images. These will show that our method is valid enough when the data have certain variety.

Keywords : Evaluation Function, Sort Data

Contact Address : **Ryoji FUKUDA**

Faculty of Engineering Oita University
700, Dan-noharu, Oita, 870-1192, JAPAN
TEL : 097-554-7860 (7860)
FAX : 097-554-7860
E-mail : rfukuda@cc.oita-u.ac.jp