



# データ解析演習 C

1. 回帰・因子分析・多次元尺度構成法

千野直仁

愛知学院大学心身科学部

於関西学院大学

2005年8月8-11日

# 目次

<b>第 1 章</b>	<b>回帰と因子分析の基礎</b>	<b>1</b>
1.1	回帰の基礎	2
1.1.1	単回帰と相関係数	2
1.1.2	重回帰と重相関	6
1.1.3	等調回帰と単調回帰	9
1.2	因子分析の基礎	11
1.2.1	サー斯顿の多因子モデル	11
<b>第 2 章</b>	<b>尺度構成の基礎</b>	<b>12</b>
2.1	自然科学における量的概念とその測定	13
2.2	心理学における量的概念とその測定	14
2.2.1	感覚尺度構成法	14
2.2.2	心理学的尺度構成法	14
2.2.3	多次元尺度構成法	17
2.3	古典的 MDS と基礎定理	18
2.3.1	ヤング・ハウスホルダーの定理	18
2.3.2	エッカート・ヤングの定理	20
2.3.3	古典的 MDS における対象間距離の推定方法と加算定数	20
2.3.4	対象間距離からの布置回復例	21
<b>第 3 章</b>	<b>多次元尺度構成法の展開</b>	<b>26</b>
3.1	トーガソン以降の主要な MDS	27
3.1.1	ノンメトリック MDS	27
3.1.2	個人差 MDS	29
3.1.3	ALSCAL	30
3.1.4	最尤 MDS	31
3.2	その他の MDS と最近の展開	33
3.2.1	多次元展開法	33
3.2.2	内積モデル	35
3.2.3	非対称 MDS	36
<b>付録 A</b>	<b>演習問題の証明</b>	<b>42</b>

# 第1章 回帰と因子分析の基礎

社会・行動科学の分野におけるデータ解析を考える時、われわれはしばしば古典的・伝統的統計解析の方法の限界を感じる。1つは、それらの方法の前提の幾つかが、社会・行動科学の分野におけるデータにはそぐわない時である。伝統的・古典的統計解析では、多くの場合定量的データを前提とし、線形回帰と相関係数、及び測定誤差の正規性が仮定される。

しかし、社会・行動科学の分野における多くの現象は、古典的・伝統的統計解析の前提となるような単純なものさしでは測定できない場合が多い。また、たとえ測定できたとしても、現象そのものは多次元적であり、そもそもそれらのものでさしにより測定された内容が、何を測っているのか、現象の背後にあると考えられる本質的な内容が何であるかも判然としない場合がある。現在では、推測統計的方法としての地位を確立した因子分析や共分散構造分析は、まさにそのような判然としない多くのものさしで測定される現象から、そこに潜在する基本的要素を推測したり、要素間の因果的關係までも推定しようとするものである。もっとも、これらの方法も相関係数という伝統的な統計解析の呪文から開放されているわけではない。

この授業の主題である多次元尺度構成法とその応用では、対象間の近さ・遠さの指標として、相関係数よりははるかにゆるい仮定のみを課す指標を仮定し、複数の対象や成員間に潜在する基本的な要素を推測したり、可能であればそれらの対象や成員間の遠近關係の変化やそのダイナミックスの予測まで行うことを意図する。

現象の予測や説明という問題も、他の分野の学問と同様、社会・行動科学の分野で重要である。統計学の分野では、そのための方法として最も伝統があるのが回帰の問題である。この授業でも、主題に入る前に、伝統的な回帰の概念と基礎的な回帰の問題をレビューし、相関の概念との関連性及び主題としての多次元尺度構成法とその応用にとってどのような関連があるのか、また最近の展開はどうなっているのかを一瞥することにする。

## 1.1 回帰の基礎

### 1.1.1 単回帰と相関係数

同一被験者に対する定量的2変数の測定値の組  $(x_1, y_1), \dots, (x_N, y_N)$  から、両変数間の直線関係の有無を検討する方法は、単純線形回帰（あるいは、単回帰）(simple linear regression) と呼ばれる。

単回帰のモデルすなわち単純線形回帰モデル (the simple linear regression model) は、

$$Y_i = \beta_0 + \beta_1 x_i + E_i, \quad (1.1)$$

と表される。ここで、 $Y_i$  は被験者  $i$  の測定値  $y_i$  を実現値 (realization) とする確率変数である。また、 $E_i$  は被験者  $i$  の測定誤差を表す確率変数である。

$E_i$  は通常、

$$E_i \sim N(0, \sigma^2), \quad (1.2)$$

すなわち、平均 0、分散  $\sigma^2$  なる正規分布 (the normal distribution) に従うと仮定される。

一方、 $\beta_0$ 、 $\beta_1$  は回帰係数 (regression coefficients) と呼ばれ、通常定数として扱われる。

(1.1) 式の大きな特徴は、説明ないし予測変数  $x_i$ ,  $i = 1, \dots, N$  が通常定数として扱われる点である。これより、確率変数  $Y_i$  は、

$$Y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2), \quad (1.3)$$

に従うと仮定される。ここで、 $Y_i$  の期待値 (expectation) は

$$E(Y_i) = \beta_0 + \beta_1 x_i. \quad (1.4)$$

(1.4) 式はしばしば

$$y = \beta_0 + \beta_1 x, \quad (1.5)$$

と書き、回帰直線 (regression line または、regression equation) と呼ばれる。(1.5) 式は、 $x$  から  $y$  の値を予測ないし説明するもので、 $y$  の  $x$  への回帰 (the regression of  $y$  on  $x$ ) と呼ばれる。

(1.5) 式の直線の方程式では、 $\beta_0$  は  $y$  軸切片であり切片パラメータ (the intercept parameter)、 $\beta_1$  は勾配を表し勾配パラメータ (the slope parameter) と呼ばれることもある。

実際のデータ  $(x_1, y_1), \dots, (x_N, y_N)$  は、必ずしも回帰直線上にはない。そこで、もし  $x$  と  $y$  の間に直線関係が予想されるのであれば、 $x_i$  に対する  $y_i$  の予測値  $\hat{y}_i$  は (1.4) 式の  $E(Y_i)$  とみなし、

$$\hat{y}_i = \beta_0 + \beta_1 x_i, \quad (1.6)$$

に対して、実測値  $y_i$  と予測値  $\hat{y}_i$  の差の2乗  $(y_i - \hat{y}_i)^2$  を被験者全体にわたり最小にするような  $\beta_0$  及び  $\beta_1$  を求めるのが、一つの方法である。

(1.1) 式より、 $Y_i = y_i$  の時、 $E_i = e_i$  と置けば、 $y_i - \hat{y}_i$  は  $e_i$  に等しいので、そのような  $\beta_0$  及び  $\beta_1$  は、

$$Q = \sum_{i=1}^N e_i^2 = \sum (y_i - \hat{y}_i)^2 = \sum (y_i - \beta_0 - \beta_1 x_i)^2, \quad (1.7)$$

を最小にするときのそれらとして求めればよい。

(1.7) 式は、未知数  $\beta_0$  及び  $\beta_1$  に関する2次関数であり、 $Q$  が最小になるための必要条件は、 $\partial Q / \partial \beta_0 = \partial Q / \partial \beta_1 = 0$  である。そこで、まず  $\partial Q / \partial \beta_0 = 0$  より、次式が得られる：

$$\beta_0 = \bar{y} - \beta_1 \bar{x}. \quad (1.8)$$

この結果を用いて、(1.7) 式を  $\beta_1$  のみの関数とみなし、 $\partial Q / \partial \beta_1 = 0$  とすることにより、次式が得られる：

$$\beta_1 = \frac{s_{xy}}{s_x^2}. \quad (1.9)$$

ここで、

$$s_{xy} = \frac{1}{N} \sum_i (x_i - \bar{x})(y_i - \bar{y}), \quad s_x^2 = \frac{1}{N} \sum_i (x_i - \bar{x})^2. \quad (1.10)$$

(1.7) 式のような誤差の2乗和を最小にするように未知数を求める方法は、一般に最小2乗法 (least squares method) と呼ばれ、それによる未知数の推定値は最小2乗推定値 (least squares estimate) と呼ばれる。以降、われわれは (1.8)、(1.9) 式の  $\beta_0$  及び  $\beta_1$  を、最小2乗推定値の意味を込めて、

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \quad \hat{\beta}_1 = \frac{s_{xy}}{s_x^2}, \quad (1.11)$$

と書くことにする。

$\hat{\beta}_0$  及び  $\hat{\beta}_1$  を用いると、(1.6) 式の  $y_i$  の推定値  $\hat{y}_i$  は、

$$\hat{y}_i = \bar{y} + \frac{s_{xy}}{s_x^2} (x_i - \bar{x}), \quad (1.12)$$

と書ける。

同一データに対して、われわれは  $x$  の  $y$  への回帰 を考えることもできる。この場合、回帰直線は

$$x = \alpha_0 + \alpha_1 y, \quad (1.13)$$

となり、同様な手続きにより、 $\alpha_0$  及び  $\alpha_1$  の最小2乗推定値  $\hat{\alpha}_0$  及び  $\hat{\alpha}_1$  を求めると、

$$\hat{\alpha}_0 = \bar{x} - \hat{\alpha}_1 \bar{y}, \quad \hat{\alpha}_1 = \frac{s_{xy}}{s_y^2}, \quad (1.14)$$

が得られる。

ここで、同一データに対する2つの回帰直線の勾配の幾何平均の2乗を計算すると、

$$\hat{\alpha}_1 \hat{\beta}_1 = \left( \sqrt{\hat{\alpha}_1 \hat{\beta}_1} \right)^2 = \left( \frac{s_{xy}}{s_x s_y} \right)^2, \quad (1.15)$$

が得られる。

(1.15) 式の右辺のカッコ内を

$$r_{xy} = \frac{s_{xy}}{s_x s_y}, \quad (1.16)$$

と書くと、 $r_{xy}$  はピアソンの偏差積率相関係数 (Pearson's product-moment correlation coefficient) になっている。

(1.12) 式を相関係数を用いて変形すると、興味深い式が得られる。すなわち、

$$\frac{s_{xy}}{s_x^2} = r_{xy} \frac{s_y}{s_x}, \quad (1.17)$$

に注意すれば、(1.12) 式は、つぎのようにも書ける：

$$\hat{y}_i - \bar{y} = r_{xy} \left( \frac{x_i - \bar{x}}{s_x} \right). \quad (1.18)$$

ここで、 $x$  が親の身長、 $y$  が子供の身長だとすると、(1.18) 式からは親の身長がその世代の身長の平均より偏差単位で大きく平均から離れていても、一般に  $|r_{xy}| < 1$  なので、その子供の身長（の予測値）は、子供の世代の身長の平均からは、その世代の身長の偏差単位で測るときには、それほど離れない、すなわち退行することがわかる。

この結果は、Guilford (1973, pp.338-339) によれば、Francis Galton が彼のいとこの Charles Darwin の進化論から示唆されて進めた遺伝の研究の中で見つけたもので、いわゆる子孫退行の法則 (the law of filial regression) と呼ばれる。

つぎに、(1.7) 式の  $Q$  が最小化された時点で得られる  $y$  や  $e$  の推定値、 $\hat{y}$  及び  $\hat{e}$  の性質を見ておこう。この時、まず (1.1) 式から

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{e}_i = \hat{y}_i + \hat{e}_i. \quad (1.19)$$

この式から容易に、

$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x} + \bar{\hat{e}}. \quad (1.20)$$

一方、 $Q$  が最小化された時点では、(1.11) 式が成り立つので、それらの式の前者を用いると、(1.20) 式から、

$$\bar{\hat{e}} = 0. \quad (1.21)$$

つぎに、(1.7) 式、(1.8) 式、及び (1.21) 式より、

$$s_{\hat{e}}^2 = \frac{1}{N} \hat{Q} = \frac{1}{N} \sum_i \left\{ (y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x}) \right\}^2, \quad (1.22)$$

が得られるので、(1.11) 式の後者を代入して整理すると、次式が得られる：

$$s_{\hat{e}}^2 = \hat{\sigma}^2 = s_y^2 - \frac{s_{xy}^2}{s_x^2} = s_y^2 (1 - r_{xy}^2). \quad (1.23)$$

ここで、(1.12) 式より、

$$\hat{y} = \frac{1}{N} \sum_i \hat{y}_i = \bar{y}, \quad (1.24)$$

に注意すると、

$$s_{\hat{y}}^2 = \frac{1}{N} \sum_i (\hat{y}_i - \hat{y})^2 = \frac{1}{N} \sum_i (\hat{y}_i - \bar{y})^2 \quad (1.25)$$

ここで再度、(1.12) 式の関係を用いると、

$$s_{\hat{y}}^2 = \frac{s_{xy}^2}{s_x^2}, \quad (1.26)$$

(1.23) 式と (1.26) 式より、

$$s_y^2 = s_{\hat{y}}^2 + s_{\varepsilon}^2, \quad \text{又は} \quad 1 = \frac{s_{\hat{y}}^2}{s_y^2} + \frac{s_{\varepsilon}^2}{s_y^2}. \quad (1.27)$$

ここで、(1.27) 式における  $s_{\hat{y}}^2/s_y^2$  は、(1.26) 式の関係を用いると、

$$\frac{s_{\hat{y}}^2}{s_y^2} = r_{xy}^2. \quad (1.28)$$

すなわち、y の分散に占める予測値  $\hat{y}$  の分散の比率は、誤差が最小化された時点では、x と y の相関係数の 2 乗に一致する。この比率は、決定係数 (coefficient of determination) とも呼ばれる。

ここで、(1.27) 式と (1.28) 式より、

$$1 - r_{xy}^2 = \frac{s_{\varepsilon}^2}{s_y^2}, \quad (1.29)$$

が成り立つので、これより

$$-1 \leq r_{xy} \leq 1, \quad (1.30)$$

が言える。もっとも、(1.30) 式は (1.16) 式の定義から、直接導くこともできる。

つぎに、 $s_{y\hat{y}}$  は定義より

$$s_{y\hat{y}} = \frac{1}{N} \sum_i (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}}) = \frac{1}{N} \sum_i (y_i - \bar{y})(\hat{y}_i - \bar{y}). \quad (1.31)$$

ここで、(1.12) 式の関係を用いると、

$$s_{y\hat{y}} = \frac{s_{xy}^2}{s_x^2}, \quad (1.32)$$

が言え、(1.26) 式に注意すると

$$s_{y\hat{y}} = s_{\hat{y}}^2. \quad (1.33)$$

そこで、この関係を用いると

$$r_{y\hat{y}} = \frac{s_{y\hat{y}}}{s_y s_{\hat{y}}} = \frac{s_{\hat{y}}}{s_y} \geq 0. \quad (1.34)$$

さらに、(1.28) 式に注意すると



$$r_{y\hat{y}} = |r_{xy}|. \quad (1.35)$$

最後に、(1.27) 式と同等な関係として

$$s_{\hat{y}\hat{e}} = 0, \quad (1.36)$$

が成り立つことを示す。これを示すには、まず

$$s_{x\hat{e}} = 0, \quad (1.37)$$

を示すとよい。

定義より

$$s_{x\hat{e}} = \frac{1}{N} \sum_i (x_i - \bar{x})(\hat{e}_i - \bar{\hat{e}}). \quad (1.38)$$

この式で、(1.21) 式及び (1.19) 式の  $\hat{e}_i$  に注意すると、

$$s_{x\hat{e}} = \frac{1}{N} \sum_i (x_i - \bar{x})(y_i - \hat{y}_i). \quad (1.39)$$

ここで、 $y_i - \hat{y}_i = (y_i - \bar{y}) - (\hat{y}_i - \bar{y})$  と置き、(1.12) 式に注意して整理すれば、(1.37) 式が得られる。

この時、

$$s_{\hat{y}\hat{e}} = \frac{1}{N} \sum_i (\hat{y}_i - \bar{\hat{y}})(\hat{e}_i - \bar{\hat{e}}) = \frac{1}{N} \sum_i (\hat{y}_i - \bar{y}) \hat{e}_i. \quad (1.40)$$

ここで、再度 (1.12) 式の関係を用いると、

$$s_{\hat{y}\hat{e}} = \frac{s_{xy}}{s_x^2} s_{x\hat{e}}, \quad (1.41)$$

が得られるので、(1.36) 式が成り立つことがわかる。

相関係数の分布については、例えば千野の以下のホームページを参照のこと：

<http://www.aichi-gakuin.ac.jp/~chino/multivar/chapter1/sec1-2-2.html>

### 1.1.2 重回帰と重相関

単回帰では、説明ないし予測変数は1つであったが、それが複数の場合、(線形)重回帰 (multiple regression) と呼ばれる。

各被験者 (i) のデータは、一般に  $(x_{i1}, \dots, x_{im}, y_i)$  となる。また、この場合のモデルは、重回帰モデル (the multiple (linear) regression model) と呼ばれ、

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_m x_{im} + E_i, \quad (1.42)$$

と表される。 $m \geq 2$  の場合、回帰係数  $\beta_0, \beta_1, \dots, \beta_m$  は、偏回帰係数 (partial regression coefficients) と呼ばれる。 $E_i$  は誤差を表す確率変数で、単回帰の場合と同様

$$E_i \sim N(0, \sigma^2), \quad (1.43)$$

が仮定される。

また、基準変数 (the criterion variable) の測定値  $y_1, \dots, y_n$  は、確率変数  $Y_1, \dots, Y_N$  の実現値とみなされる。

例えば、 $m = 2$  の時、回帰方程式は、

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2, \quad (1.44)$$

の形を取り、3次元ユークリッド空間内の平面の方程式を表す。誤差なしデータの場合、 $N$ 組のデータ点  $(x_{i1}, x_{i2}, y_i)$ ,  $i = 1, \dots, N$  は、すべてこの平面上に位置することになる。

実際のデータでは、誤差はつきものなので、われわれは単回帰の場合と同様

$$Q = \sum_{i=1}^N e_i^2 = \sum (y_i - \hat{y}_i)^2 = \sum (y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_m x_{im})^2, \quad (1.45)$$

を最小にするような方程式の係数  $\beta_1, \beta_2, \dots, \beta_m$  を求める。単回帰の場合の (1.7) 式は、この式の特別なケースである。

そのような係数は、

$$\frac{\partial Q}{\partial \beta_0} = \frac{\partial Q}{\partial \beta_1} = \dots = \frac{\partial Q}{\partial \beta_m} = 0, \quad (1.46)$$

を満たす必要があるので、単回帰の場合と同様、まず  $\partial Q / \partial \beta_0 = 0$  より、

$$\beta_0 = \bar{y} - \sum_{j=1}^m \beta_j \bar{x}_j. \quad (1.47)$$

が得られる。ここで、 $\bar{x}_j$  は、第  $j$  説明ないしは予測変数の測定値の平均であり、 $\bar{y}$  は基準変数  $Y$  の実測値、すなわち  $y$  の平均である。単回帰の場合の (1.8) 式は、(1.47) 式の特別なケースであることは明らかである。

(1.47) 式を用いると、(1.45) 式は

$$Q = \sum_{i=1}^N \left[ (y_i - \bar{y}) - \sum_{j=1}^m \beta_j (x_{ij} - \bar{x}_j) \right]^2, \quad (1.48)$$

と書ける。

ここで、

$$\mathbf{X}_m = \begin{pmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \dots & x_{1m} - \bar{x}_m \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \dots & x_{2m} - \bar{x}_m \\ \dots & \dots & \dots & \dots \\ x_{N1} - \bar{x}_1 & x_{N2} - \bar{x}_2 & \dots & x_{Nm} - \bar{x}_m \end{pmatrix}, \quad (1.49)$$

を定義すると、 $Q$  を最小にする  $\beta_1, \beta_2, \dots, \beta_m$  は、

$$\boldsymbol{\beta}_m = (\beta_1, \beta_2, \dots, \beta_m)^t, \quad (1.50)$$

$$\mathbf{y}_B = (y_1 - \bar{y}, y_2 - \bar{y}, \dots, y_N - \bar{y})^t, \quad (1.51)$$

と書けば、

$$\boldsymbol{\beta}_m = (\mathbf{X}_m^t \mathbf{X}_m)^{-1} \mathbf{X}_m^t \mathbf{y}_B, \quad (1.52)$$

と書ける。(1.52) 式は、 $\beta_0$  を除く偏回帰係数が説明変数の全測定値の組と基準変数の実測値の組が得られれば、それらを用いて計算する式を与える。単回帰の場合の (1.9) 式は、(1.52) 式の特別な場合になっている。

いずれにせよ、(1.47) 式及び (1.52) 式より、われわれは  $N$  組のデータから重回帰モデルの未知数の一部である  $\beta_0, \beta_1, \dots, \beta_m$  の最小 2 乗推定値、 $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_m$  を

$$\hat{\beta}_0 = \bar{y} - \sum_{j=1}^m \hat{\beta}_j \bar{x}_j. \quad (1.53)$$

$$\hat{\beta}_m = (\hat{\beta}_1, \dots, \hat{\beta}_m)^t = (\mathbf{X}_m^t \mathbf{X}_m)^{-1} \mathbf{X}_m^t \mathbf{y}_B, \quad (1.54)$$

として求めればよい。

(1.45) 式の  $Q$  が最小化された時点では、単回帰と同様 (1.27) 式が成り立つ。ただし、この場合、予測値 (の推定値)  $\hat{y}$  及び誤差の推定値  $\hat{e}$  は、(1.42) 式に注意して、

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \dots + \hat{\beta}_m x_{im} + \hat{e}_i, \quad (1.55)$$

なる関係を満たすものとする。

一方、単回帰における (1.28) 式にあたるものは、重回帰では

$$\frac{s_{\hat{y}}^2}{s_y^2} = R^2, \quad (1.56)$$

となる。この式の右辺の  $R^2$  は、基準変数の分散に占める予測値の分散であり、重相関係数の 2 乗 (squared multiple correlation coefficient) と呼ばれ、しばしば SMC と略記される。

単回帰の場合の (1.34) 式と同一の関係が重回帰の場合も成り立つことが証明できるので、(1.56) 式の関係とあわせると、

$$\frac{s_{\hat{y}}^2}{s_y^2} = R^2 = r_{y\hat{y}}^2. \quad (1.57)$$

実は、重相関係数  $R$  は、もともと  $y$  の実測値とその予測値との相関係数として定義される。すなわち、

$$R = r_{y\hat{y}}, \quad (1.58)$$

が成り立つ。そこで、(1.34) 式及び (1.58) 式から

$$0 \leq R \leq 1, \quad (1.59)$$

であることがわかる。

(1.54) 式により得られる偏回帰係数の推定値を用いると、各被験者 ( $i$ ) の基準変数の実測値  $y_i$  は、(1.55) 式のように書ける。そこで、(1.53) 式の  $\hat{\beta}_0$  を (1.55) 式に代入して整理すれば、

$$y_i - \bar{y} = \sum_{j=1}^m \hat{\beta}_j (x_{ij} - \bar{x}_j) + \hat{e}_i, \quad (1.60)$$

と書けるので、この式の両辺を  $s_y$  で割り、整理すると

$$z_{y_i} = \sum_{j=1}^m \hat{\beta}_j^* z_{x_{ij}} + \tilde{e}_i, \quad (1.61)$$

と書ける。ここで、

$$z_{y_i} = \frac{y_i - \bar{y}}{s_y}, \quad z_{x_{ij}} = \frac{x_{ij} - \bar{x}_j}{s_j}, \quad \tilde{e}_i = \frac{\hat{e}_i}{s_y}, \quad (1.62)$$

であり、

$$\hat{\beta}_j^* = \hat{\beta}_j \left( \frac{s_j}{s_y} \right), \quad (1.63)$$

とする。うへの  $\hat{\beta}_j^*$  は、偏回帰係数の推定値を当該説明ないしは予測変数の標準偏差と、基準変数の標準偏差で調節したもので、標準偏回帰係数 (the standard partial regression coefficients) と呼ばれる。逆に、標準化されていない (もとの) 偏回帰係数は、このような調節がなされていないので、説明ないしは予測変数間の基準変数に対する効きの強さの比較を直接これらの値によって行うことはできない。

(1.54) 式を、(1.63) 式の関係を用いて変形すると、最終的には次式が得られる：

$$\hat{\beta}_m^* = (\hat{\beta}_1^*, \dots, \hat{\beta}_m^*)^t = \mathbf{R}_m^{-1} \mathbf{r}_y, \quad (1.64)$$

ここで、 $\mathbf{r}_y = (r_{x_1 y}, \dots, r_{x_m y})^t$ 、すなわち各説明ないし予測変数と基準変数の相関係数を並べたものである。この式は  $\mathbf{r}_y = \mathbf{R}_m \hat{\beta}_m^*$  とも書けることに注意して、これをスカラー表現すれば、

$$r_{jy} = r_{x_j y} = \sum_{k=1}^m \hat{\beta}_k^* r_{jk} = \hat{\beta}_j^* + \sum_{k \neq j} \hat{\beta}_k^* r_{jk}, \quad (1.65)$$

と書ける。このことは、基準変数  $y$  への説明ないし予測変数  $x_j$  の効果が、 $x_j$  (正確には、 $z_{x_j}$ ) による直接的効果と、 $x_j$  (正確には、 $z_{x_j}$ ) の変化に伴う  $x_k$  (正確には、 $z_{x_k}$ ) ( $k = 1, \dots, m$  で  $k \neq j$ ) の変化を通じての間接的効果  $\hat{\beta}_k^* r_{jk}$  ( $k = 1, \dots, m$  で  $k \neq j$ ) の合計  $m$  個の和として表されることを意味する。そこで、標準偏回帰係数は、パス係数あるいは経路係数 (path coefficient) と呼ばれることもある。

重相関係数の検定については、例えば千野の下記のホームページを参照のこと：

<http://www.aichi-gakuin.ac.jp/~chino/multivar/chapter2/sec2-3.html>

また、偏回帰係数の検定については、例えば千野の下記のホームページを参照のこと：

<http://www.aichi-gakuin.ac.jp/~chino/multivar/chapter2/sec2-4.html>

### 1.1.3 等調回帰と単調回帰

単回帰の場合、(1.1) 式の形の線形モデルをデータに適合させるが、社会・行動科学の分野では、説明ないし予測変数が順序情報しか得られない場合も多い。このような場合、(1.5) 式の形の陽な線形方程式をデータに当てはめることには無理がある。

そのような場合に、当てはめるべき方程式  $f(x)$  の形は特定せず、つぎのようにその性質のみを定義し、一種の最小2乗法によりデータを  $f(x)$  に当てはめる回帰分析もあり、等調回帰 (isotonic regression) と呼ばれる (例えば、Robertson, Wright, & Dykstra, 1988)：

#### 定義1 関数の等調性

$X$  は有限集合  $x_1, x_2, \dots, x_k$  で、単純順序 (simple order)  $x_1 \preceq x_2 \preceq \dots \preceq x_k$  を持つとする。 $X$  上の関数  $f$  は、もし、

$$f(x_1) \leq f(x_2) \leq \dots \leq f(x_k), \quad (1.66)$$

ならば、等調 (isotonic) という。

ここで、単純順序とはつぎのように定義される：

定義2  $X$  上の二項関係 (a binary relation)  $\preceq$  は、以下の条件を満たすならば  $X$  上の単純順序である：

1. 反射的： $x \in X$  に対して、 $x \preceq x$ 、

2. 推移的:  $x, y, z \in X$ ,  $x \preceq y$ , かつ  $y \preceq z$  ならば,  $x \preceq z$ ,
3. 反対称的:  $x, y \in X$ ,  $x \preceq y$  かつ  $y \preceq x$  ならば,  $x = y$ ,
4.  $X$  のどの2つの要素も比較可能:  $x, y \in X$  ならば,  $x \preceq y$  または  $y \preceq x$

### 定義3 等調回帰

$g$  は  $X$  上の所与の関数とする。 $X$  上の関数  $g^*$  は  $g^*$  が等調の時に限り

$$\sum_{x \in X} [g(x) - f(x)]^2 w(x), \quad (1.67)$$

を最小化するならば、重み  $w$  に関する  $g$  の等調回帰 (an isotonic regression of  $g$  with weight  $w$ ) という。isotonic regression は、単調回帰 (monotonic regression) と訳されることもあるが、当てはめるべき関数  $f(x)$  が非減少 (nondecreasing) の時等調、非増大 (nonincreasing) の時反調 (または逆調) (antitonic)、いずれの場合も単調 (monotonic) と呼ばれるので、正確には等調回帰と訳するのがよからう。

等調回帰の考え方は、計量心理学の分野では、今回の講義の主題である多次元尺度構成法 (multidimensional scaling、略して MDS) の分野で非計量 MDS の一つの方法として Kruskal (1964a, b) が利用している。Kruskal は、彼の非計量 MDS のアルゴリズムを、単調回帰の問題として位置づけている。

(註1) 相関係数の概念は、伝統的な多変量解析の中核をなす概念であるが、回帰の概念はそれより一世紀近く古い。McCullagh & Nelder (1989, p.9) によれば、既に19世紀の初頭(1805年)に Legendre が古典的線形モデル (classical linear model) の範疇で、現代流に言えば重回帰分析モデルの偏回帰係数を最小2乗推定する方法を提案している。また、誤差項の正規性は Gauss が1809年に導入している。彼らは天文学のデータに対して、これらの方法を適用した。

(註2) 無相関と(統計的)独立とは同一ではないので注意が必要である。両者の関係については、まず確率変数の期待値の定義を行い、2つの確率変数の積の期待値  $E(XY)$  が、2つの事象の独立性を用いて、 $X$  と  $Y$  が独立の時、 $E(XY) = E(X)E(Y)$  を導く。つぎに、2つの確率変数の共分散  $Cov(X, Y)$  の定義式から、「 $X$  と  $Y$  が独立の時共分散はゼロ、すなわち無相関である」ことを導けばよい。この逆は、必ずしも成り立たないので注意が必要である。

(註3) (註1)でも見たように、回帰の考え方及び伝統的な線形回帰分析の歴史は長いが、この節で紹介した重回帰モデルの基準変数は1変量 (univariate) に過ぎない。これに対して、その後、基準変数が多変量の場合の多変量重回帰分析 (別名、多変量線形モデル (the multivariate linear model) (例えば、Muirhead, 1982 参照)、当てはめるモデルが線形ではなく非線形の場合の(一変量)非線形回帰分析 (univariate nonlinear regression)・多変量非線形回帰分析 (multivariate nonlinear regression) (例えば、Gallant, 1987 参照) が考案されている。実際、例えば SAS の NLIN プロシジャは、前者の一変量非線形回帰分析を実行できる。さらに最近では、最初 Nelder & Wedderburn (1972) により導入された一般化線形モデル (the generalized linear model) やその拡張モデル等が、反復測定デザインデータの解析がらみで数多く提案されている。一般化線形モデルでは、古典的な線形回帰分析もその前提である、モデルの直接的な線形性や誤差分布の正規性の仮定はもはや必要ではない。最近の拡張された一般化線形モデルは、基準変数や説明ないし予測変数に反復測定変数を許すばかりではなく、基準変数も説明ないし予測変数とも連続変数以外の変数も許容する大変汎用なものになっている。例えば、SAS の GENMOD プロシジャは、このような幅広いデータに対応する非線形モデルである。これらも含めた、最近の反復測定デザインデータの分析方法については、例えば千野 (2003) を参照されたい。これについては、国立情報学研究所の以下の検索サイトで、筆者の名前を入力すれば、その先頭論文として pdf ファイルが提供されている：

<http://ci.nii.ac.jp/cinii/servlet/CiNiiTop#>

## 1.2 因子分析の基礎

### 1.2.1 サーストンの多因子モデル

因子分析は、もともと心理学者が中心となり知能の分析のために開発されたが、いまや心理学に限らず、社会・行動科学、医学等の分野でも幅広く利用されている、基礎的な統計解析の方法となっている。古典的モデルとしての計量心理学者 Thurstone (1931) が提案した多因子模型はつぎのようである：

$$Z_{ij} = a_{j1}F_{i1} + a_{j2}F_{i2} + \dots + a_{jr}F_{ir} + b_jG_{ij}. \quad (1.68)$$

ここで、 $Z_{ij}$  は、第  $i$  サンプルの第  $j$  変数の値で、平均 0、分散 1 に標準化されたもの、 $a_{j1}, \dots, a_{jr}, b_j$  は、因子の係数又は負荷量 (loadings) と呼ばれる未知数、 $F_{i1}, F_{i2}, \dots, F_{ir}, G_{ij}$  は、第  $i$  サンプルの因子の値又は因子得点 (factor scores) と呼ばれる。

因子分析の基礎的知識については、例えば以下の千野の以下のホームページを参照されたい：

<http://www.aichi-gakuin.ac.jp/~chino//multivar/chapter3/sec3-1-1.html>

伝統的な多因子モデルに代表される因子分析は、現在では探索的因子分析 (exploratory factor analysis) として位置づけられており、確証的因子分析 (confirmatory factor analysis) の延長上に位置づけられる共分散構造分析 (analysis of covariance structures) あるいは別名 構造方程式モデル (structural equation models) がよく用いられる。これについても、例えば千野の下記のホームページを参照されたい：

<http://www.aichi-gakuin.ac.jp/~chino//multivar/chapter5/sec5-1.html>

なお、多変量解析・多変量データ解析全般については柳井 (1994) を、多変量解析の方法を射影子の方法を用いて統一的に議論したものとしては竹内・柳井 (1972) を推薦する。

## 第2章 尺度構成の基礎

この章では、多次元尺度構成法について議論するに先立ち、まず自然科学における量的概念とその測定について簡単に紹介し、つぎに心理学における量的概念とその測定が心理測定の研究の歴史の中でどのようにして議論され発展してきたかについて簡単に紹介する。最後にいわゆる古典的多次元尺度構成法（古典的 MDS）とその基本定理について紹介する。

自然科学における量的測定に関しては、著名な論理実証主義哲学者の一人であるカルナップ (Carnap, 1966) の著書「物理学の哲学的基礎」の一節を紹介し、自然科学的測定の客観性・厳密さとその特徴に言及する。とりわけ物理学的測定では、尺度構成は主観を排し、さらに対象の持つ何らかの性質や当該状態の生起するメカニズムに関する理論に基づいてなされることがわかる。

自然科学的測定に対して心理学的測定においては、尺度構成すなわち複数の対象や刺激の次元もしくは多次元尺度上への位置づけに際しては、ほとんどの場合それらの持つ性質や当該状態の生起するメカニズムに関して理論が確立されていないと言ってよい。例えば、社会的ないしは心理的对象に対する態度を測定しようとするとき、対象に対する態度自身がどのようなまたどれだけの次元から成り立っているのか、さらには対象に対する被験者の態度はどのような脳内のメカニズムが働き最終的な態度の表出が行われるのか等については、ほとんどわかっていないのが実情である。

また、心理学的測定においては、対象の位置づけが被験者またはグループにより異なる場合も多々ある。この点も、心理学的尺度構成法が自然科学的尺度構成法と比べて大きく異なる点であろう。このような事情を反映して、心理学における尺度構成法の歴史はまず物理学的対象の持つ量的属性が被験者が感じた心理学的値とどのような対応関係にあるのかとの問題設定からスタートしている。いわゆる感覚尺度構成法と呼ばれており、19世紀中ごろから精神物理学的測定において議論された。

これに対して、対象の測定を必ずしもその物理的特性との関連なしで行えることを20世紀前半に計量心理学者のサーストンらが示し、現在では一次的尺度構成法と位置づけられる心理学的尺度構成法の研究が発展した。心理学的尺度構成法では、対象の位置づけは対象の持つ物理的性質や状態の生起するメカニズムから離れ、対象に対する心理学的判断の誤差分布等の仮定から対象を次元連続体上に位置づける、すなわち次元尺度の構成を行う方法が発展した。

(一次的) 尺度構成法はリチャードソン、トーガソンらにより50年代には多次元尺度構成法 (MDS) に発展した。初期の MDS は、現在では古典的 MDS と呼ばれ、対象間の(非)類似度データは原則的には比率尺度レベルで測定されている必要があったが、その制約も60年代以降シェパード・クラスカルらにより取り払われ、順序尺度レベルで測定された対象間の(非)類似度データから対象を多次元距離空間に位置づけるノンメトリック MDS、対象に対する判断の個人差を考慮する個人差 MDS、誤差分布に正規分布等を仮定しモデルの未知パラメータを最尤法で推定する最尤 MDS などが多くの計量心理学者により提案された。この章では、それらの MDS の基本と成る古典的 MDS の基本定理について最後にふれる。

## 2.1 自然科学における量的概念とその測定

論理実証主義哲学者として著名な Carnap (1966) は、著書「物理学の哲学的基礎」の中で、科学上の3つの主要概念として、分類概念 (classificatory concept)、比較概念 (comparative concept)、及び量的概念 (quantitative concept) をあげている。彼の主張を要約・補足するとつぎのようである：

「分類概念」は、対象（や状態）以降必要な場合のみ両者を使い分けるが、それに言及しない場合には、対象もしくは状態を意味するものとする。Carnap の原文では、「状態」の方は使われていないが、この資料ではこれを追加した）をあるクラスに割り付ける概念であり、生物学における魚類、は虫類、ほ乳類などがその代表例である。また対象を「重い」とか「軽い」と分類するだけであれば、やはりこれらは分類概念に過ぎない。

一方、2つの対象を比較して「より重い」とか「より軽い」や、対象の状態を「より暖かい」とか「より冷たい」とか言う場合、これらは「比較概念」となる。比較概念の場合、それにより複数の対象や状態を比較するための客観的経験的手続きが必要となる。例えば、重さの概念の場合、そのためには天秤と2つの規則 – 2つの対象  $a$ 、 $b$  の相等性 (equivalence) に関するものおよび大小関係に関するもの（ここではより小さい  $less\ than$  という関係とし） – が必要である。これら2つの経験的關係 (empirical relations) を、以降  $E_M(a, b)$  及び  $L_M(a, b)$  と表記するものとする。

比較概念で重要な点は、それが自然の事実に適用されるかどうかという問題とは別に 関係の論理的構造 に縛られる、という点である。つまり、ここで3つの対象や状態を  $a$ 、 $b$ 、 $c$ 、として、関係  $E$  は対称的 (symmetric) すなわち  $E(a, b) \Rightarrow E(b, a)$  であり、かつ推移的 (transitive) すなわち  $E(a, b)$  かつ  $E(b, c) \Rightarrow E(a, c)$  でなければならない。一方、関係  $L$  は非対称的 (asymmetric) すなわち  $L(a, b) \Rightarrow L(b, a)$  でないし、推移的すなわち  $L(a, b)$  かつ  $L(b, c) \Rightarrow L(a, c)$  が成り立たなければならない。

「量的概念」は 対象が数値を持つ概念 であるが、この値に達するための客観的手続きとその規則が必要である。そのためには、一般にある量  $M$  について、つぎの5つの規則が必要であるという：

1.  $E_M(a, b) \Rightarrow M(a) = M(b)$
2.  $L_M(a, b) \Rightarrow M(a) < M(b)$
3. 特定の値  $s$ （通常ゼロ）を特定の対象や状態  $p$  に割り付ける ( $M(p) = s$ )。

例えば、温度の場合、凝固しつつある水の状態  $p$  に対してゼロなる値を割り付ける ( $M(p) = 0$ )。

4. 単位の規則、と呼ばれるもので、もう1つの容易に認識され再現できる対象や状態  $q$  を明確に定義し、この量の第2の特定の値  $t$  を割り付ける ( $M(q) = t$ )。通常は1である。

例えば摂氏目盛りでは、 $t=100$  を沸騰状態の水  $q$  に対して割り付ける ( $M(q) = 100$ )。

5. 量の差の相等性の条件についての規則

$$ED_M(a, b, c, d) \Rightarrow M(a) - M(b) = M(c) - M(d)$$

ここで、 $ED_M(a, b, c, d)$  は経験的条件とする。例えば、温度の場合、温度計に用いられる物質の体積について、水銀のどれか2つの体積の差が、他の体積との差に等しいとき、目盛りが温度の等しい差を示すような温度計を作らなければならない。そのような温度計ができれば、われわれは上記  $ED_M(a, b, c, d)$  をつぎのように表現することができる： $ED_M(a, b, c, d) \equiv d(a, b) = d(c, d)$

ここで、 $d(a, b)$  は温度計が物体  $a$  に接触している時の目盛りと物体  $b$  に接触している時の目盛りとの差を表すものとする。

これまでの議論で注意すべきは、Carnap の議論している重さや温かさ（温度）の測定は対象や状態の心理的判断や主観的判断に関するものではなく、自然科学的・客観的判断に関するものである、という点である。言い換えれば、



物理学的測定では、尺度構成は物理学的対象の持つ何らかの性質や当該状態の生起するメカニズムについての理論に基づいて行われる、

と言う点が重要である。

## 2.2 心理学における量的概念とその測定

### 2.2.1 感覚尺度構成法

心理学における量的概念とその測定に関する理論は、歴史的にはライプチヒ大学の物理学教授 G.T. Fechner (1801-1887) により創始された 精神物理学 (psychophysics) にさかのぼることができよう。彼のめざしたのは精神と身体の関係に関する精密理論であるが、具体的には物理学的刺激強度と心理学的感覚強度との数学的関数関係の究明であり、よく知られたフェヒナーの法則 (Fechner's law)  $R = k \log(S/b)$  をウエーバーの法則から理論的に導いた (Fechner, 1860)。ここで、 $R$  は、感覚強度であり、 $S$  は刺激強度、 $b$  は閾値すなわち感覚が生じかつ消失する刺激値の単位である。つまり、人が感じる感覚の大きさは、(物理) 刺激の絶対的大きさではなく、刺激の大きさの対数に比例するというわけである。

彼の法則が、基本的には被験者に2つの物理刺激の感覚強度の差(の有無)の判断をさせることにより間接的に感覚量を測定するのに対して、スティーブンス (Stevens, S. S, 1951) は被験者に物理的な標準刺激の感覚量に対する比較刺激の直接的な感覚量の(比の)推定を行わせることにより、著名なスティーブンスの法則 ないしは ベキ法則  $R = aS^l$  を提案した。

これらの尺度構成法は、前節で述べた物理学における量的測定とは異なり、(原則的には) 物理学的対象やその状態のもつ何らかの量との対応で、いわゆる 心理学的連続体 (psychological continuum) 上に各対象や状態を位置づけることにより、心理学的尺度、より限定的には感覚尺度 (sensory scale) を構成するので、感覚尺度構成法と呼ばれることもある。いずれにせよ、

この種の尺度では、対象やそれが持つ何らかの状態そのものの性質や状態の生起するメカニズムについての理論から尺度を構成するわけではない

点に注意が必要である。

心理学的尺度構成を、精神物理学的測定、とりわけ物理刺激の弁別過程の考察から、さらに発展させるきっかけを作ったのが、つぎの節で述べるサーストン (Thurstone, L. L, 1887-1955) の幾つかの仕事である。彼は、心理学的尺度を構成する対象や状態は必ずしも物理的对象や状態である必要はなく、物理学的刺激や状態、心理学的刺激や状態における測定誤差や判断誤差に関する統計学的分布の情報が特定できれば、それらの対象や状態を心理学的連続体上に位置づけることができることを示した。

### 2.2.2 心理学的尺度構成法

前節の感覚尺度構成法も広義には心理学的尺度構成法と呼ぶことができるが、ここでは後者を前者と区別することにする。前者の構成では、物理学的刺激との対応で心理学的尺度を構成するが、心理学的対象や状態の場合、物理学の場合のそれらと異なり、例えば知能、学力、態度、物理的刺激に対する心理的判断などのように、対象やそれが持つ何らかの状態そのものの性質や、状態の生起するメカニズムについての理論が確立していない場合がほとんどと言えよう。そのような場合の幾つかについて、旧来の精神物理学的方法とは全く異なる方法により(心理学的) 尺度を構成する方法を最初に提案したのが、サーストンである。それらは、絶対尺度 (absolute scale)、比較判断の法則 (the law of comparative judgement)、及び カテゴリ

リー判断の法則 (the law of categorical judgement) であり、以下に統計学辞典 (1989) に従って簡単に紹介する。

まず、絶対尺度 (Thurstone, 1925, 1928; Thurstone & Ackerson, 1929) は、ビネー式知能検査のような複数の項目から成る心理学的検査を複数の年齢や学年の生徒に実施して得られる正答率のデータから、個々の検査項目を1つの尺度(これを彼は絶対尺度と呼ぶ)上に位置づけ、さらには各年齢や学年の生徒のグループも同一尺度上に位置づける方法である。ここで、1) 能力や学力の分布は正規分布に従う、2) 各年齢段階の(母)標準偏差は必ずしも等しくない、と仮定する。例えば、年齢  $A_i$  及び年齢  $A_j$  のグループの得点の絶対尺度上での分布が、それぞれ  $N(\mu_i, \sigma_i^2)$  及び  $N(\mu_j, \sigma_j^2)$ 、検査項目  $T_k$  の絶対尺度上での位置を  $t_k$  とする。この時、 $Z_{ik} = (t_k - \mu_i)/\sigma_i$ 、同様に  $Z_{jk} = (t_k - \mu_j)/\sigma_j$  を定義すれば、つぎの関係が成り立つ：

$$Z_{ik} = (\sigma_j/\sigma_i) Z_{jk} + (\mu_j - \mu_i)/\sigma_i. \quad (2.1)$$

したがって、観測された年齢グループ  $A_i$  及び  $A_j$  の項目  $T_k$  に対する観測された正答率  $\hat{p}_{ik}$  及び  $\hat{p}_{jk}$  から、正規偏差  $Z_{ik}$  及び  $Z_{jk}$  を推定すれば、(2.1) 式を直線と見なした場合の勾配  $(\sigma_j/\sigma_i)$  及び  $y$ -切片  $(\mu_j - \mu_i)/\sigma_i$  をデータ  $(\hat{Z}_{ik}, \hat{Z}_{jk}, k = 1, 2, \dots, n)$  から推定できる。ここで、 $n$  はグループ  $A_i$  及び  $A_j$  の双方でデータが得られている項目数である。そうすれば、例えば最低年齢のグループの分布の平均  $\mu_1$  と標準偏差  $\sigma_1$  をそれぞれゼロ、及び1と決めればすべての年齢グループの分布の平均と標準偏差を推定できる。詳細は、Thurstone (1925) を参照のこと。また、これにより得られた  $\mu_j$  及び  $\sigma_j$  ( $j = 1, 2, \dots$ ) の推定値から、テストによっては、尺度値の絶対ゼロ (absolute zero) を決定できる。この議論については Thurstone (1928; Thurstone & Ackerson, 1929) を参照のこと。

比較判断の法則 (Thurstone, 1927a, b) は、サーストンの精神物理学的測定に対する考察から生まれたもので、刺激対象についての(1次元的)尺度構成を以下のように物理学的刺激との対応を仮定することなしに行おうとするものである。彼は、一対比較法による被験者の対刺激  $O_j$ 、 $O_k$  間の優劣、軽重、好悪などの判断に対して、何らかの(一次元的)心理学的連続体上で弁別過程 (discriminal process)  $v_j$ 、 $v_k$  が生じ、それらはつぎのように書けると仮定する：

$$v_j = \mu_j + \varepsilon_j, \quad \varepsilon_j \sim N(0, \sigma_j^2), \quad v_k = \mu_k + \varepsilon_k, \quad \varepsilon_k \sim N(0, \sigma_k^2)$$

ここで、 $\mu$  は尺度値、 $\varepsilon$  は誤差項である。

これらを用いて、つぎの弁別差異 (discriminal difference)  $u_{jk}$  が定義できる：

$$u_{jk} = v_j - v_k = \mu_{jk} + \varepsilon_{jk}. \quad (2.2)$$

ここで、 $\mu_{jk} = \mu_j - \mu_k$ 、 $\varepsilon_{jk} \sim N(0, \sigma_{j-k}^2)$ 、 $\sigma_{j-k}^2 = \sigma_j^2 + \sigma_k^2 - 2\rho_{jk}\sigma_j\sigma_k$ 。また、 $\rho_{jk}$  は  $\varepsilon_j$  と  $\varepsilon_k$  との相関係数である。

この時、 $O_j$  が  $O_k$  より優れている(重い、好ましいなど)と判断される確率は、

$$P_{jk} = \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma_{j-k}} \exp\left\{-\frac{1}{2}\left(\frac{u - \mu_{jk}}{\sigma_{j-k}}\right)^2\right\} du = \Phi(Z_{jk}). \quad (2.3)$$

ここで、 $Z_{jk} = \mu_{jk}/\sigma_{j-k}$  とする。また、 $\Phi$  は単位正規分布の分布関数である。一般に  $Z_{jk}$  は、いわゆる正規偏差 (normal deviate) あるいはより正確には正規同値偏差 (normal equivalent deviate) と呼ばれる。

[演習 1.1] (2.3) 式を証明せよ。

$Z_{jk}$  と  $\sigma_{j-k}^2$  の関係から、次式のいわゆるサーストンの比較判断の法則が導かれる：

$$\mu_j - \mu_k = Z_{jk} \sqrt{\sigma_j^2 + \sigma_k^2 - 2\rho_{jk}\sigma_j\sigma_k}. \quad (2.4)$$

(2.4) 式は、一般には未知数が多すぎて解けないが、例えばサーストンの Case V の仮定 ( $\rho_{jk} = 0$ ,  $\sigma_j^2 = \sigma_k^2 = \text{constant}$ ,  $j, k = 1, 2, \dots, N$ ) に加え、 $\sum_{k=1}^N \mu_k = 0$ 、及び ( $\sigma_{j-k} = 1$ ) なる仮定をすれば、 $\hat{\mu}_j^0 = \frac{1}{N} \sum_{k=1}^N Z_{jk}$ ,  $j = 1, 2, \dots, N$  として解ける。なお、 $Z_{jk}$  は (2.3) 式から  $\hat{Z}_{jk} = z_{jk} = \Phi^{-1}(\hat{P}_{jk})$ ,  $\hat{P}_{jk} = \sum_{i=1}^{N_{jk}} s_{jki}/N_{jk}$  として推定する。ここで、 $N_{jk}$  は (j,k) 対の判断についてのサンプル数であり、 $s_{jki}$  は対象  $O_j$  が対象  $O_k$  より優れている (重い、好ましい等) と判断される時 1、そうでない時 0 とするものとする。

なお、上の解は統計的推測を行うには不十分で、これを行うには最小ノーマットカイ二乗法 (Bock & Jones, 1968) や最尤法 (Arbuckle & Nugent, 1973) が用いられる。また、サーストンモデルに対して、ブラッドリー (Bradley, R. A.) らの BTFL モデル (Bradley-Terry-Ford-Luce model) では、 $P_{jk} = s_j/(s_j + s_k)$  が仮定される。ここで、 $s_j$  は対象  $O_j$  に対する尺度値である。この時、最尤解は  $s_j = \sum_k N_{jk} p_{jk} / \{\sum_k N_{jk} / (s_j + s_k)\}$  となる (Hohle, 1966)。

カテゴリ判断の法則は、サーストンが系列間隔法 (method of successive intervals) により得られたデータに対して比較判断の法則を応用したものを、トーガソンが発展させたものである (Bock & Jones, 1968; Saffir, 1937; Torgerson, 1958)。被験者は、 $n$  個の対象のそれぞれを  $m+1$  個のカテゴリからなる評定尺度上のいずれかのカテゴリに該当するとして評定する。このようなカテゴリ判断の場合、次の 2 つの弁別過程が生ずると仮定する：

$$v_j = \mu_j + \varepsilon_j, \quad \varepsilon_j \sim N(0, \delta_j^2), \quad v_k = \tau_k + \varepsilon_k, \quad \varepsilon_k \sim N(0, \gamma_k^2)$$

ここで、 $\mu_j$  は対象  $O_j$  の尺度値、 $\tau_k$  は弁別連続体上の第  $k$  カテゴリと第  $k+1$  カテゴリの境界点の値である。また、 $\varepsilon_j$  と  $\varepsilon_k$  の同時分布は 2 変量正規分布に従うとする。この時、弁別差異

$$v_{jk} = v_j - v_k = \mu_j - \tau_k + (\varepsilon_j - \varepsilon_k), \quad (2.5)$$

は正規分布に従い、平均と分散はつぎのようになる：

$$E(v_{jk}) = \mu_j - \tau_k, \quad V(v_{jk}) = \sigma_{j-k}^2 = \delta_j^2 + \gamma_k^2 - 2\rho_{jk}\delta_j\gamma_k.$$

これらより、被験者が対象  $O_j$  を評定尺度の第  $k$  カテゴリ以下として評定する確率  $P_{jk}$  は、つぎのように書ける：

$$P_{jk} = \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}\sigma_{j-k}} \exp\left\{-\frac{1}{2}\left(\frac{v - (\mu_j - \tau_k)}{\sigma_{j-k}}\right)^2\right\} dv = \Phi(Z_{jk}). \quad (2.6)$$

ここで、 $Z_{jk} = (\tau_k - \mu_j)/\sigma_{j-k}$  とする。

[演習 1.2] (2.6) 式を証明せよ。

この場合も、一般にはデータ数  $mn$  に対して、未知パラメータ数が多すぎて解けない。Torgerson (1958) は、これらの未知パラメータにいろいろな制約を課した場合のパラメータの求め方について述べている。また、適当な初期値を与え、最小ノーマットカイ二乗解 (Bock & Jones, 1968) や最尤解 (Schönemann & Tucker, 1967) を求めることもできる。

この節で述べた心理学的尺度構成法では、もはや (心理学的) 対象の持つ性質や当該状態の生起するメカニズムから離れ、均質な集団内での得点の分布や潜在的な一次元尺度上での反応の統計学的分布の仮定のみから、尺度構成がなされる、

点に注意が必要である。

ここで取り上げた尺度構成法以外にも、心理学の領域では例えば態度測定分野で提案されてきたボガードスの社会的距離尺度 (social distance scale) (Bogardus, 1925)、サーストンの等現間隔法 (method of equal-appearing intervals) (Thurstone, 1931)、リッカートの集積評定法 (summated rating method) (Likert, 1932)、ガットマンの尺度分析 (法) あるいは尺度解析 (法) (scale analysis, scalogram analysis) (Guttman, 1944)、クームズの展開法 (unfolding) (Coombs, 1950, 1964) などがあるが、ここでは後に紹介する展開法を発展させた諸手法を除き、多次元尺度構成法を次節で述べるような狭義の意味で用いるので取り上げない。

なお、これらのうち、ガットマンの尺度解析 (法) は、その後単なる尺度の解析法にとどまらず、ガットマンとその後継者達によって発展してきた ファセット理論 (Facet theory) 中のファセット・アナリシス (Facet analysis) の1つの方法として位置づけられている (例えば、Borg & Shye, 1995; 木村他, 2002; Manabe, 2001)。

### 2.2.3 多次元尺度構成法

これまで紹介してきた尺度構成法では、対象はすべて一次元連続体 (unidimensional continuum) 上に位置づけらるという暗黙の仮定がなされている。しかし、社会行動科学とりわけ心理学が扱う対象や状態では、必ずしもこの仮定は成り立たないと見るほうが自然であろう。心理学的尺度構成法で扱う対象や状態は、客観的な対象や状態そのものではなく、それらに対する人の認知的、感情的、あるいは行動的成分であるので、サーストンが指摘しているように (Thurstone, 1927b)、必ずしも物理的大きさやサイズを持つものに限らない。したがって、それらの内包する連続体の次元数、さらにはそれにより構成される空間構造も心理学の現状のレベルでは判然としない。ただし、心理学や大脳生理学の進展によっては、例えば対象に対する好き嫌いは脳内物質を直接客観的に簡単な装置により測定できる日が来れば、1.1 節で述べた温度の測定のような尺度構成が、可能になるかも知れない。

いずれにせよ、現状ではその次元数や次元の内容が判然としない対象 (や状態) を、それらの間の類似度ないしは非類似度 (判断) データをもとに、一般的には次元数及び空間構造が特定された多次元連続体上に位置づける方法が、多次元尺度構成法 (multidimensional scaling、略して MDS) である。ただし、MDS では、類似度ないし非類似度データをもとに、対象に (多次元の) 数値を付与する類似の幾つかの方法に比べて、対象を位置づける或いは埋め込む空間の数学的構造を厳密に定義し、類似度ないし非類似度データにフィットさせる点に大きな特徴がある。もちろん、類似度ないし非類似度データをもとに対象に数値を付与する方法を MDS に含める、いわゆる広義の定義も存在するが、ここでは以下のような狭義の意味での定義を、MDS とする。以下の定義は高根 (1980) による：

MDS は対象間の (非) 類似性の程度を示す測度が与えられたとき、対象を多次元空間内の点として表し、点間の距離が観測された (非) 類似性と最も良く一致するように点の布置を定める方法である。

一方、千野 (1997) は、この節で述べる伝統的な MDS を含む一般的な非対称 MDS (asymmetric MDS) を、距離空間の概念をより明確にする形で、つぎのように定義している：

非対称 MDS とは、非対称 (非) 類似度行列をもとに、対象間の類似度が何らかの距離空間上の距離の関数であるという仮定のもとに、観測結果としての (非) 類似度データの情報をできるだけ満たすように、各対象を多次元距離空間内の点として位置づける方法である。

いずれにせよ、MDS の定義からは、観測された (非) 類似性データに対して如何なる距離空間 (metricspace)

を仮定するかが重要課題となる。MDS の分野では通常いわゆるユークリッド空間 (Euclidean space) が、より一般的にはつぎのミンコフスキーの  $r$  メトリック (Minkowski's  $r$ -metric) が仮定される：

$$d_{jk} = \left[ \sum_{a=1}^t |x_{ja} - x_{ka}|^r \right]^{1/r}, \quad (r \geq 1.0), \quad (2.7)$$

$r$  メトリックは、数学の分野では  $l_p$  ノルム ( $l_p$ -norm) として知られている。

ミンコフスキーの  $r$  メトリックで、例えば  $r = 1$  のときは、市街化計量 (city block metric) またはマンハッタン計量 (Manhattan metric) と呼ばれる。また、 $r = 2$  のときは、よく知られたユークリッド計量となる。

MDS は、リチャードソン (Richardson, 1938) のアイデアを、ヤング・ハウスホルダー (Young & Householder, 1938) やトーガソン (Torgerson, 1952) らが発展させたもので、現在では古典的尺度構成法 (classical scaling) とか計量的 MDS (metric MDS) などと呼ばれる。計量的 MDS では、あとで見るように複数の対象間の非類似度が比率尺度、とりわけ対象間のユークリッド距離として推定されている場合に、次節で述べるヤング・ハウスホルダーの定理 (Young-Householder's theorem) (Young & Householder, 1938) をもとに、対象を多次元ユークリッド空間の点として位置づける。

## 2.3 古典的 MDS と基礎定理

### 2.3.1 ヤング・ハウスホルダーの定理

ヤング・ハウスホルダーの定理は、複数の対象間のユークリッド距離が既知の場合、対象が埋め込まれる空間の次元数 (dimensionality) や、それらのユークリッド距離がユークリッド空間上の真の点間距離であるための必要十分条件などについて述べたものであり、MDS の基礎定理と見ることができる。

ここで、 $n$  個の対象  $O_1, O_2, \dots, O_n$  のうち、 $O_n$  を原点とし、残りの  $n-1$  個の対象を終点とする (列) ベクトルを  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}$  とする。さらに、対象  $O_i$  の第  $j$  軸の座標値を  $x_{ij}$  とし、行列  $\mathbf{X} = \{x_{ij}\}$  を定義しよう。行列  $\mathbf{X}$  は、 $n-1$  個の対象の埋め込まれる空間の次元数分の座標値を対象ごと各行に並べたものである。

この時、この空間の次元数  $r$  は、

$$\mathbf{B} = \{b_{ij}\} = \mathbf{X}\mathbf{X}^t, \quad (2.8)$$

と置けば、 $r = \text{rank } \mathbf{X} = \text{rank } \mathbf{B}$  となる。行列  $\mathbf{B}$  は、 $n-1$  個の対象間の内積 (inner product) を要素とする行列で、内積行列 (inner product matrix) である。ここで、(2.8) 式の行列  $\mathbf{B}$  の要素は、

$$b_{ij} = \mathbf{x}_i^t \mathbf{x}_j = x_{i1}x_{j1} + x_{i2}x_{j2} + \dots + x_{ir}x_{jr}, \quad (2.9)$$

つまり、一般に座標値が既知ならば内積は簡単に (2.9) 式から計算できるが、MDS では対象間の非類似度データから各対象の座標値を求める必要があり、この式は使えない。

つぎに、ベクトル  $\mathbf{v}_{ij} = \mathbf{x}_j - \mathbf{x}_i$  を定義してみよう。この時、このベクトルの長さ (対象  $O_i$  と  $O_j$  間のユークリッド距離  $d_{ij}$ ) をユークリッドノルム (the Euclidean norm) を用いて表現すれば、

$$\|\mathbf{v}_{ij}\|^2 = d_{ij}^2 = \mathbf{v}_{ij}^t \mathbf{v}_{ij} = (\mathbf{x}_j - \mathbf{x}_i)^t (\mathbf{x}_j - \mathbf{x}_i) = \|\mathbf{x}_i\|^2 + \|\mathbf{x}_j\|^2 - 2b_{ij}. \quad (2.10)$$

と書ける。これは、よく知られた余弦法則 (the cosine law) の一つの表現である。これより、

$$b_{ij} = \frac{1}{2}(d_{in}^2 + d_{jn}^2 - d_{ij}^2). \quad (2.11)$$

この式は、対象間の距離が既知ならば、対象間の内積が求まることを示しており、これにより計算された内積行列から対象の座標値を復元しようとするのが、古典的 MDS の基本的な発想である。

いずれにせよ、これより対象が埋め込まれる空間の次元数については、

点間距離  $d_{ij}$  が既知の (原点を  $O_n$  とする) 複数の点の集合 (組) の次元数は、(2.11) 式で定義される  $b_{ij}$  を要素に持つ  $n-1$  次の正方行列  $B$  の階数に等しい。

定理では、さらに上記点間距離の組がユークリッド空間上の真の点間距離であるための条件をつぎのように示している：

定理1 (Young-Householder)  $d_{ij} = d_{ji}$  の組がユークリッド空間の真の点の集合 (組) 間の相互距離であるための必要十分条件は、(2.11) 式で定義される要素からなる (内積) 行列  $B$  が正の半定符号 (positive semi-definite) であることである。

[演習 1.3] ヤング・ハウスホルダーの定理を証明せよ。

実際、 $B$  が正の半定符号ならば、行列  $B$  の固有値 (eigenvalue) はすべて正かゼロであり、われわれはこの行列を

$$B = OD^2O^t = (OD)(OD)^t = AA^t, \quad (2.12)$$

と固有値分解 (eigenvalue-eigenvector decomposition) すればよい。ここで、行列  $D^2$  は対角行列で、対角要素には行列  $B$  の固有値が並ぶ。(2.12) 式は  $B = XX^t$  に注意すれば、 $X = USV^t$  と特異値分解 (singular value decomposition) するのと同様である。なお、固有値分解については、Wilkinson (1965)、Lancaster & Tismenetsky (1985) が詳しい。一方、特異値分解については、Lawson & Hanson (1974) 及び Stewart & Sun (1990) が参考になる。

ヤング・ハウスホルダーの論文の脚注を見ると、彼らの論文は (MDS のアイデアを最初に提案した) リチャードソン (M. W. Richardson)、及びガリクセン (H. Gulliksen) の示唆に答えるために書かれた、とあり、MDS に対するリチャードソンの貢献の大きさがしのばれる。

ところで、行列  $B$  は対象  $O_n$  を原点とした場合の内積行列であり、どの対象を原点に取るかで行列  $A$  は異なる、すなわち対象の座標値は異なるが、対象間距離に誤差が含まれない場合は、対象の配置は原点と軸の回転を除き不変である。ただし、現実のデータの場合、対象間距離の推定値には誤差が含まれるので、特定の対象を原点に選ぶことは好ましくない。これを避ける1つの方法は、原点を対象の重心にとることである。これを行うには、対象間距離 (の推定値) を要素とする  $n$  行  $n$  列の行列  $T = \{d_{ij}^2\}$  が得られたら、つぎに示す中心化行列 (centering matrix) (Horst, 1965):

$$G_n = I_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^t, \quad (2.13)$$

を用いて、 $n$  個の対象の重心を原点とするつぎのような内積行列を計算する:

$$B_c = -\frac{1}{2} G_n T G_n^t. \quad (2.14)$$

ここで、行列  $G_n$  及び  $B_c$  は、共に  $n$  行  $n$  列の行列であり、 $\mathbf{1}_n$  は、1 ばかりが  $n$  個並んだ列ベクトルである。また、 $I_n$  は  $n$  次の単位行列 (unit matrix) もしくは恒等行列 (identity matrix) である。さらに、 $G_n$  は具体的には、つぎのように書ける：

$$G_n = \begin{pmatrix} 1 - \frac{1}{n} & -\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} & -\frac{1}{n} \\ -\frac{1}{n} & 1 - \frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} & -\frac{1}{n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -\frac{1}{n} & -\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} & 1 - \frac{1}{n} \end{pmatrix}. \quad (2.15)$$

さて、行列  $B_c$  が計算できたならば、あとはこれをつぎのように分解すれば、

$$B_c = X_c X_c^t, \quad (2.16)$$

得られた  $n$  行  $r$  列の行列  $X_c$  は、 $B_c$  が正の半定符号であれば、 $n$  個の対象の重心を原点とする各対象の  $r$  次元ユークリッド空間内の座標値を要素とする行列と見做せる。

### 2.3.2 エッカート・ヤングの定理

ヤング・ハウスホルダーの定理と並び MDS の基礎定理の 1 つとみなされる定理が上記特異値分解、とりわけエッカート・ヤングの定理 (Eckart & Young, 1936) である。特異値分解は、19 世紀の後半に発見されたものであるが、その後この分解は多くの研究者により再発見されたり、拡張されている (Stewart & Sun, 1990)。現在では、特異値分解は複素行列の場合にも拡張されているが、ここでは実数行列の場合についてのみ紹介する。ここで、行列  $A$  を  $m$  行  $n$  列の階数  $r$  の行列とする。この時、行列  $A$  は、

$$A = USV^t, \quad (2.17)$$

と分解でき、特異値分解と呼ばれる。ここで、行列  $U$  及び  $V$  はそれぞれ  $m$  次及び  $n$  次の直交行列 (orthogonal matrix) である。また、行列  $S$  は  $m$  行  $n$  列の対角行列で、非負の値を取り、特異値と呼ばれる。これらのうち、 $k$  個は正である。ここで、通常対角行列は正方行列であるが、この場合は矩形行列である点に注意が必要である。なお、ここでの分解は最大階数 (full rank) 分解である。

エッカート・ヤング (Eckart & Young, 1936) は、特異値分解を正方行列から矩形行列の場合に一般化し、かつ任意の矩形行列をそれより階数の低い他の行列で近似した場合の階数  $r$  のベストフィット行列を求める問題に利用し、ベストフィット行列はその矩形行列の特異値分解を用いて得られることを示した。

特異値分解は、上に述べたように古典的 MDS でも固有値分解の変わりに使えるし、後に簡単に紹介する個人差 MDS の 1 つであるタッカー・メシクの PVA (the Point of View Analysis) (Tucker & Messick, 1963) でも用いられている重要な定理である。

### 2.3.3 古典的 MDS における対象間距離の推定方法と加算定数

ヤング・ハウスホルダーの定理から、われわれは複数の対象間のユークリッド距離がわかれば対象の多次元空間上への位置づけ、すなわち各対象の座標値を求めることができることはわかったが、それではそのような対象相互の距離をどのようにしたら推定することができるであろうか。例えば、トーガソン (Torgerson, 1958) は、つぎのような方法を提案している。

彼は、まずマンセルの色立体を構成する色刺激の中から色相はすべて同じ赤で明度と彩度の異なる 9 刺激を選び出し、被験者に完全三組法 (complete method of triads) により刺激間の類似度判断 (similarity judgment) を求めた。ここで、完全三組法とは複数の刺激対象間の類似度を 2 対間の比較ではなく 3 対間で行う方法で、3 つの刺激を  $i, j, k$  とすれば、被験者には例えば「刺激  $i$  は刺激  $j$  と刺激  $k$  のどちらにより似ているか」と尋ねる。この種の判断を複数の被験者に対して行うことにより、われわれは「刺激  $i$  が刺激  $k$  に比べてより刺激  $j$  に似ている」と判断された比率  ${}_i\hat{P}_{jk}$  を観測できる。一般に、このような 3 対の選び方は、刺激数を  $n$  とすれば  $\frac{1}{2}n(n-1)(n-2)$  通りある。

トーガソンの古典的 MDS では、これらの比率を手にしたとして、つぎにすべての刺激間の距離の差

$${}_i d_{jk} = d_{ik} - d_{ji}, \quad (2.18)$$

を推定する。

つまり、われわれはここではまず完全三組法により得られた「刺激  $i$  が刺激  $k$  に比べてより刺激  $j$  に似ている」という判断比率データから、刺激を埋め込むべき何らかの心理学的空間内での刺激間の距離  $d_{ik}$  と  $d_{ji}$  の差を推定しようというわけである。

そのために、トーガソンは、2.2.2 節でふれたサーストンの比較判断の法則を上述のデータに対して適用することを考えた。比較判断の法則を完全三組法により得られる刺激間の類似度判断比率  ${}_i\hat{P}_{jk}$  の問題に適用すると、サーストンのいう Case V 条件のもとでは、 ${}_iP_{jk}$  と (2.18) 式で与えられる刺激間距離の差  ${}_i d_{jk}$  との間には次の関係が成り立つ：

$${}_iP_{jk} = \int_{-\infty}^{{}_i d_{jk}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx. \quad (2.19)$$

(2.19) 式における左辺の比率は理論比率であり、一方上述の完全三組法により得られる比率は観測比率  ${}_i\hat{P}_{jk}$  でしかないので、われわれが (2.16) 式により手にするのは、(2.18) 式の左辺の  ${}_i d_{jk}$  そのものではなくその推定値  ${}_i\hat{d}_{jk}$  に過ぎないことに注意したい。

いずれにせよ、このようにしてわれわれが刺激間距離の差の推定値  ${}_i\hat{d}_{jk}$  を手にしたならば、それらの推定値から刺激間の真の距離を推定する必要がある。Torgerson は、これを (2.18) 式により得られるすべての推定値  ${}_i\hat{d}_{jk}$  を用いて最小 2 乗法 (least squares procedure) により求めることを提案した。

ただし、この方法により得られる刺激間距離の推定値  ${}_i\hat{d}_{jk}$  は、必ずしも正確な距離空間の特性を備えていない。というのは、この方法により得られる距離の推定値は必ずしも絶対原点 (absolute origin) からの距離ではないのである。Torgerson は、このようにして得られる距離を、絶対原点からの距離と区別して、比較距離 (comparative distance) と呼んでいる。

一般に距離空間の特性を備えている距離  $d_{jk}$  と比較距離  $h_{jk}$  との間にはつぎの関係が成り立つ：

$$d_{jk} = h_{jk} + c. \quad (2.20)$$

ここで、定数  $c$  を Torgerson は加算定数 (additive constant) と呼んだ。

最も簡単な加算定数推定の方法は、トーガソン (Torgerson, 1952) によるもので、つぎのようである。ここで、真の距離  $d_{jk}$  についてはユークリッド空間の場合、もし 3 対象が直線上にあれば、 $d_{ik} = d_{ij} + d_{jk}$  が成り立つので、この式と (2.20) 式から  $c = h_{ik} - h_{ij} - h_{jk}$  とすればよい。実際にはそのような 3 点あるとは限らない。そこで、(直線上にない) 3 対象間の距離について三角不等式が成り立つことを利用し、これに (2.20) 式を適用し整理すると、 $c \geq h_{ik} - h_{ij} - h_{jk}$  が成り立つので、すべての 3 つの刺激組について  $h_{ik} - h_{ij} - h_{jk}$  の最大値  $c_1$  を計算し、 $c_1$  を加算定数の下限とみなす。すなわち

$$c \geq c_1 = \max (h_{ik} - h_{ij} - h_{jk}), \quad (2.21)$$

さらに、(2.20) 式から対象間の真の距離  $d_{jk}$  が負にならないとの制約を課すと、すべての 2 つの刺激に関する比較距離  $h_{jk}$  の符号を反転したものの最大値を  $c_2$  (すなわち、 $c_2 = \max(-h_{jk})$ ) として、

$$c = \max (c_1, c_2), \quad (2.22)$$

とすればよい。その後、加算定数を求める新たな方法が提案されている (例えば、Messick & Abelson, 1956; Saito, 1978)。

### 2.3.4 対象間距離からの布置回復例

#### (a) 架空データ例

ここでは、前節の (2.19) 式の意味での真の距離  $d_{jk}$  の二乗値が、表 1.1 のようになったと仮定する。この場合、4 つの対象を頂点とする図形はひし形をなす。このデータすなわち対象間距離から、各対象の座標



表 2.1: 4つの刺激間の距離の推定値の2乗の架空の例

色	1	2	3	4
1	0	5/4	1	5/4
2	5/4	0	5/4	4
3	1	5/4	0	5/4
4	5/4	4	5/4	0

値を計算し（布置の回復）、布置から逆算した対象間距離の二乗値が表 1.1 の距離の二乗値に一致することを示そう。この表の行列を  $T$  と見なして、前節の行列  $B_c$  を計算すると、

$$2B_c = \begin{pmatrix} .50 & .00 & -.50 & .00 \\ .00 & 2.00 & .00 & -2.00 \\ -.50 & .00 & .50 & .00 \\ .00 & -2.00 & .00 & 2.00 \end{pmatrix},$$

となる。

つぎに上の  $2B_c$  を固有値分解すると、固有値は 4.0 及び 1.0 で残り 2 つはゼロとなる。したがって、非ゼロ固有値を対角要素に並べた行列  $\Lambda$  は、

$$\Lambda = \begin{pmatrix} 4.0 & 0.0 \\ 0.0 & 1.0 \end{pmatrix},$$

一方、これらの固有値に対応する固有ベクトルを列に持つ行列  $Y_c$  は、

$$Y_c = \begin{pmatrix} 0.00000 & 0.70711 \\ 0.70711 & 0.00000 \\ 0.00000 & -0.70711 \\ -0.70711 & 0.00000 \end{pmatrix}.$$

これらより、 $2B_c$  を (2.16) 式の右辺のような形にするには、 $Y_c$  の第 1 行のすべての要素には第 1 固有値の平方根 2.0 を、同じく第 2 行のすべての要素には第 2 固有値の平方根 1.0 を掛ければよい。この操作は、行列表現では  $Y_c \Lambda^{\frac{1}{2}}$  と書けるので、

$$Y_c \Lambda^{\frac{1}{2}} = \begin{pmatrix} 0.00000 & 0.70711 \\ 1.41422 & 0.00000 \\ 0.00000 & -0.70711 \\ -1.41422 & 0.00000 \end{pmatrix},$$

となる。上の行列の各行が表 1.1 の 2 乗距離行列  $T$  から (2.16) 式を用いて計算された 4 つの刺激の布置である。明らかにこの布置は、布置座標全体を  $1/1.41422$  倍し、4 つの色の間の点間ユークリッド距離の二乗を計算すると、表 1.1 のそれらに完全に一致する。

#### (b) 都市間距離の例

つぎの例は、Kruskal & Wish (1978) にある 10 都市間の飛行マイル数から都市についての地図を再構成する例である。このデータは、SAS の MDS プロシジャに関するマニュアルの中の最初の例としてプログラムつきであげられており、ここではそれをそのまま以下に引用する (SAS Institute, 1999) :

```

data city; title 'Analysis of Flying Mileages Between Ten U.S. cities';
input (atlanta chicago denver houston losangeles miami newyork
      sanfran seattle washdc) (5.) @56 city $15.;
datalines;
  0 Atlanta
  587 0 Chicago
  1212 920 0 Denver
  701 940 879 0 Houston
  1936 1745 831 1374 0 Los Angeles
  604 1188 1726 968 2339 0 Miami
  748 713 1631 1420 2451 1092 0 New York
  2139 1858 949 1645 347 2594 2571 0 San Francisco
  2182 1737 1021 1891 959 2734 2408 678 0 Seattle
  543 597 1494 1220 2300 923 205 2442 2329 0 Washington D.C.
;
proc mds data=city level=absolute out=out;
  id city;
run;

run;

```

最後のマクロは、グラフィックオプションがインストールされている場合には、きれいな刺激布置を出力するが、インストールされていない場合は、このマクロの代わりに print プロシジャで一時ファイル out の内容を出力させると、布置座標のみはつぎのように出力する：

Analysis of Flying Mileages Between Ten U.S. cities						
Badness- of-Fit Change in Convergence						
Iteration	Type	Criterion	Criterion	Measure		
0	Initial	0.003273	.	0.8562		
1	Lev-Mar	0.001689	0.001584	0.005128		
Convergence criterion is satisfied.						
OBS	_DIMENS_	_TYPE_	city	_NAME_	Dim1	Dim2
1	2	CRITERION			0.00	.
2	2	CONFIG	Atlanta	atlanta	717.82	142.251
3	2	CONFIG	Chicago	chicago	382.27	-341.504
4	2	CONFIG	Denver	denver	-481.62	-25.515
5	2	CONFIG	Houston	houston	161.74	572.424
6	2	CONFIG	Los Angeles	losangeles	-1203.57	382.927
7	2	CONFIG	Miami	miami	1132.65	578.403
8	2	CONFIG	New York	newyork	1072.89	-515.601
9	2	CONFIG	San Francisco	sanfran	-1419.84	108.687
10	2	CONFIG	Seattle	seattle	-1342.01	-568.788
11	2	CONFIG	Washington D.C.	washdc	979.65	-333.285

なお、上記結果は SAS の出力の一部を削除したものである。また、plotit マクロによるプロット結果は、図 1.1 に示す。

図 1.1 Kruskal & Wish (1978) の都市間のマイル数データの 2 次元布置 (当日提示)

### (c) マンセルの色立体の例

最後の例は、マンセル色立体 (Munsell Color Solid) に対する MDS の応用例である。CIE (Commission Internationale de l'Eclairage, 国際照明委員会) による CIE 表色系が光源や物体からの色刺激の特性を測色的な手続きにより表すのに対して、マンセル表色系では周知のように物体表面の色の見え (知覚色) を色の 3 属性 (色相 (hue)、明度 (value)、彩度 (chroma)) に従って 3 次元空間の点として表す (大山他, 1994)。

マンセル色立体では、色相についてはまず基本となる赤 (R)、黄 (Y)、緑 (G)、青 (B)、紫 (P) の 5 色が円周上に等間隔に配置され、つぎにそれぞれの 2 分点に中間色が配置され合計 10 分割されている。さらにそれぞれは知覚的に等間隔となるように調整され 100 分割されて環状に配置されているので、色相環と呼ばれる。一方、明度に関しては理想の黒を 0、理想の白を 10 としてその間が知覚的に等間隔になるよう等分されている。また、彩度については無彩色を 0、最も彩度の高い色票を 10 として等分されている。

したがって、明度を一定とするときの色相と彩度の異なる色は、平面状の同心円と中心からの放射状の直

線の交点に位置づけられる。ここで、同心円の半径の長さが彩度の大きさに対応する。

トーガソン (1951) は、色相 (赤) を一定として明度と彩度を変えた 9 色に対して、被験者に完全三つ組み法で色票間の類似度判断を求め、彼の提案した古典的 MDS により 9 色のマンセル色立体上の布置が再現されるかどうかを検討した (Torgerson, 1958)。MDS により得られた 9 色の布置 (図 1.2) は、マンセル色立体上のそれらの布置と近いものとなった。

図 1.2 Torgerson (1951) のマンセル色立体上の 9 色の 2 次元布置 (当日提示)

マンセル色立体については、その後多くの研究者が古典的 MDS 等を用いており、とりわけ印東は、自らの実験も含めたトーガソン以来のマンセル色に関する 19 の実験データの検討結果から興味深い推論を行っている (Indow, 1988)。彼は、トーガソンの方法をリーマン空間の場合に拡張した DMRPD (Direct Mapping in Riemannian space through Power-transformed Distance) を提案し (Indow, 1983)、図 1.3 のような興味深い結果を得ている。この図は、Indow & Aoki による、明度を一定とした 178 色の色刺激の DMRPD による 2 次元ユークリッド布置である (Indow, 1988)。ここでの stress 値は、次節で述べるクラスカルのノンメトリック MDS におけるモデルのフィットの良さの指標と形式的には等しいが、クラスカル流のいわゆる単調回帰法を適用したわけではない (例えば、Indow & Watanabe, 1988)。

図 1.3 Indow-Aoki によるマンセル色立体上の色刺激の 2 次元ユークリッド布置 (当日提示)

図 1.3 が、視空間としてユークリッド空間を仮定した場合の布置であるのに対して、図 1.4 はリーマン空間 (とりわけ曲率が正の場合) を仮定した場合の布置で、データの当てはまりは必ずしもベストではないが、色刺激の布置はよりマンセル色立体上の布置に近づいていることがわかる。

図 1.4 Indow-Aoki によるマンセル色立体上の色刺激の正の曲率を持つリーマン空間を仮定した場合の 2 次元布置 (当日提示)

この結果は、ルーンブルクらの提案した两眼視空間のリーマン空間モデル (例えば、Blank, 1953; Luneburg, 1950) に対する興味深いデータを提供する。

## 第3章 多次元尺度構成法の展開

第2章では、尺度構成の基礎について述べた。この章では、そこで紹介した尺度構成法の基礎に基づき60年代以降計量心理学の分野で提案されてきた主要な多次元尺度構成法を取り上げ、それらの方法がどのようにして複数の対象を多次元尺度上に位置づけるかについて、基本的な考え方についてのみ述べる。多次元尺度構成法は既に60年以上の研究の歴史があり、もちろんここでそれらすべてについて紹介することは不可能であるので、それらについては、高根(1980)、斉藤(1980)、Cox & Cox(2001)などを参照されたい。

最初に取り上げるのは、ノンメトリック MDS である。古典的 MDS では、最終的に手にすべき対象間の(非)類似度データは、原則として比率尺度レベルで得られていないといけませんが、この制約をはずしデータは順序尺度レベルで得られていることを前提とした MDS が何人かの研究者により60年代から提案され出した。ここでは、それらのうち、クラスカルの MDSCAL、ガットマン・リンゴーの SSA を紹介する。ノンメトリック MDS では、観測データは順序データであるが、もちろん対象は多次元距離空間内の点として位置づけられることに注意が必要である。

つぎに個人差 MDS について取り上げる。心理学等の社会行動科学の領域では、対象に対する認知や感情に個人差がつきものである。この点を考慮するのが60年代から70年代にかけて提案された個人差 MDS である。個人差 MDS では、対象に対する多次元配置は対象を認知する個人によりゆがむと仮定し、何らかの方法で対象の配置のゆがみを表現する。ここでは、代表的な個人差 MDS としてタッカーとメッシュックの PVA と、キャロルとチャンの INDSICAL を取り上げる。

70年代の後半には、高根らが古典的 MDS、ノンメトリック MDS、及び個人差 MDS などを1つの統一したアルゴリズム上で実現する方法 ALSCAL を提案した(Takane, Young, & De Leeuw, 1977)。ALSCAL では、データのタイプは比率尺度、間隔尺度、順序尺度、名義尺度を仮定し、各種類似度・非類似度データに対する MDS が可能である。現在では、例えば国際的な統計ソフトの1つである SAS にも MDS プロシージャの中で、ほぼ実現できる。

70年代から80年代にかけては、それまでの最小二乗的方法による MDS にかわり、誤差に何らかの分布を仮定し、最尤法を用いて対象の配置等の未知パラメータを推定する最尤 MDS がラムジーや高根らにより整備されてきた。ここではそれらのうち、ラムジーによる MULTISCALE と高根による MAXSCAL について紹介する。最尤 MDS を用いると、各対象の配置座標の(漸近的)信頼領域を求めたり、尤度比検定や AIC による配置の次元数の検討も可能になる。もっとも、そのためには誤差分布の仮定が正しくなければならない。

古典的 MDS 以降の上記の主要な MDS の他にも、興味深い方法が幾つかある。ここではそれらのうち、多次元展開法、内積モデルを紹介する。まず、展開法は通常の MDS と異なりデータは対象間の(非)類似度データを正方形行列にまとめたものではなく、複数刺激や対象に対する複数個人の選考データから、刺激・対象及び個人を同一の多次元距離空間上に位置づける方法であり、データは矩形行列にまとめられる。展開法は、最初クームズにより50年代に始められたもので、一次元順序尺度構成の方法としてスタートしたが、その後60年代から70年代にかけて非計量的展開法や計量的展開法へと発展したものである。ここでは、最初のクームズの展開法の考え方と、シェーネマンによる計量的展開法についてのみ紹介する。

従来の MDS では、対象間の(非)類似度に対して多次元距離空間上の距離を当てはめるモデルが主流であったが、対象相互の距離の比推定結果から対象間の内積を直接推定する内積モデルも存在する。ここでは、エックマンの多次元(率)尺度構成法について簡単に紹介する。最後に、MDS の最近の展開の1つとして非対称 MDS に若干言及する。

### 3.1 トーガソン以降の主要な MDS

#### 3.1.1 ノンメトリック MDS

##### MDSCAL

古典的 MDS では、データとしての対象相互の非類似度に比率尺度が仮定されているが、心理学の分野ではそのようなデータはきわめて稀である。これに対して、1960 年代から 1970 年代にかけて、対象相互の（非）類似度が順序尺度レベルで測定されているデータに対する MDS が何人かの計量心理学者により開発された (Guttman, 1968; Kruskal, 1964a, 1964b; Lingoes, 1973; Shepard, 1962a, 1962b)。それらは非計量的 MDS と呼ばれている。例えば、Kruskal の MDSCAL では、つぎのストレス (stress) という量

$$S = \sqrt{\frac{\sum_{j < k} \sum_{j < k} (d_{jk} - \hat{d}_{jk})^2}{\sum_{j < k} \sum_{j < k} d_{jk}^2}}, \quad (3.1)$$

を定義し、 $S$  を最小にするような対象の布置を求める。ここで、 $\hat{d}_{jk}$  は、この値が観測された対象  $j$  から対象  $k$  への非類似度  $\delta_{jk}$  と同順位になるという条件下で、 $S$  を最小にする数値である。

ストレスの大きさは、データのモデルに対する適合度の指標であるが、クラスカルは原論文 (1964a) で、ストレスがそれぞれ 0.2、0.1、0.05、0.025、0.0 に対して、（非類似度と距離の間の適合度は）悪い (poor)、かなりよい (fair)、良い (good)、すばらしい (excellent)、完全 (perfect) と呼んでいる。

一方、クラスカルが最初に提案した (3.1) 式のストレスは、現在ではストレス 1 (STRESS1) と呼ばれ、つぎのストレス 2 (STRESS2) と区別される：

$$S_2 = \sqrt{\frac{\sum_{j < k} \sum_{j < k} (d_{jk} - \hat{d}_{jk})^2}{\sum_{j < k} \sum_{j < k} (d_{jk} - \bar{d}_{\bullet\bullet})^2}},$$

(2.1) 式だと、すべての距離  $d_{jk} = c$  の場合  $d_{jk} = \hat{d}_{jk} = c$  とでき、ストレスはゼロと成る。つまり、データがモデルに完全に適合することになる。ストレス 2 (Kruskal and Carroll, 1969) は、距離が均一化する方向の解に対してブレーキをかけるような分母を定義していることになる。

より正確には、MDSCAL では原則的には対象間の距離  $d_{jk}$  が観測された非類似性の値  $\delta_{jk}$  と可能な限り弱単調 (weak monotonicity) な関係、すなわち

$\delta_{jk}$  が  $\delta_{lm}$  より小ならば、 $d_{jk}$  は  $d_{lm}$  以下となる、

ように、対象の布置を定める。しかし、観測値には誤差がつきものであるので、これを完全に満たすような布置が存在するとは限らない。そこで、（真の）対象間距離に誤差が加わった  $\hat{d}_{jk}$ 、すなわち  $\hat{d}_{jk} = d_{jk} + \varepsilon_{jk}$  を考え、この  $\hat{d}_{jk}$  が観測される非類似性の値  $\delta_{jk}$  と厳密な意味で単調関係にあると考える。つまり、

$\delta_{jk}$  が  $\delta_{lm}$  より小ならば、 $\hat{d}_{jk}$  は  $\hat{d}_{lm}$  以下となる、

この時、上記の誤差  $\varepsilon_{jk}$  は単調性からのずれと見なせる。(3.1) 式のストレスは、単調性からのずれをデータ全体で考えるものである。これを最小にするような方法は、最小二乗単調回帰法 (least squares monotonic regression) と呼ばれている。

実は、Kruskal の単調回帰法はいわゆる等調回帰法 (isotonic regression) の特別なケースであることを、Robertson, et al. (1988) が指摘している。彼らの表現を使うとすれば、“単調” という言葉には、非減少

(nondecreasing) な関係を意味する”等調”と、非増大 (nonincreasing) を意味する”反調” (antitonic) の両方の意味が含まれる。もっとも、Kruskal (1964b, p.115) も、monotonic regression をそのままではなく、”monotonic regression” と断っている点は興味深い。

MDSCAL では、対象を位置づけるべき距離空間としては、市街化距離やユークリッド空間を含む上述のミンコフスキーの  $r$  メトリックの定義された空間を仮定する。

図 2.1 は、Indow (1988) に掲載されている Indow & Ohsumi (1972) の 60 色のマンセル色立体上の色刺激の MDSCAL による 2 次元布置である。

図 2.1 Indow & Ohsumi (1972) のマンセル色立体上の 60 色の 2 次元布置 (当日提示)

また、図 2.2 は、この布置での  $d_{jk}$  と  $\hat{d}_{jk}$  のプロット結果であり、やはり Indow (1988) に掲載されているものである。

図 2.2 Indow & Ohsumi (1972) の 2 次元布置における  $d_{jk}$  と  $\hat{d}_{jk}$  のプロット (当日提示)

## SSA

Kruskal とほぼ同時期に彼とは異なる方法で、順位データに対する MDS、SSA を開発したのが Guttman と Lingoes である。SSA では、MDSCAL の  $r$  メトリックをも含む準計量 (semi-metric)

$$d_{jk} = d[(x_{j1} - x_{k1})^2, (x_{j2} - x_{k2})^2, \dots, (x_{jr} - x_{kr})^2], \quad (3.2)$$

を持つ空間を仮定する。

SSA では、つぎの単調性係数 (coefficient of monotonicity)

$$\mu = \frac{\sum_{j < k} e_{jk} \hat{D}_{jk} d_{jk}}{\sqrt{\sum_{j < k} e_{jk} \hat{D}_{jk}^2 \sum_{j < k} e_{jk} d_{jk}^2}}, \quad (3.3)$$

を最大にするような対象の布置を求める。ここで、 $D_{jk}$  は観測された非類似度  $\delta_{jk}$  の順位に関する情報  $D_{jk}$  (SSA では、これを距離順位数 distance ranking number と呼ぶ) であり、 $\hat{D}_{jk}$  は  $D_{jk}$  の順位を反映すると制約づけられた量である。また、 $e_{jk}$  は  $\delta_{jk}$  が欠測値であれば 0、存在すれば 1 なる量である。 $\mu$  を最大化することは、 $K = \sqrt{1 - \mu^2}$  を最小化することに等しい。 $K$  は離間係数 (coefficient of alienation) と呼ばれ、近年のプログラムではこちらが使われている。

SSA では、タイの取り扱いに関連して、3 種類の単調性を区別している。1 つは、準 (強) 単調性 (semi-strong monotonicity)、2 つ目は 強単調性 (strong monotonicity)、3 つ目は 弱単調性 (weak monotonicity) である。

SSA には、SSA-I、SSA-II、などデータの性質に応じてたくさんのプログラムが用意されており、Lingoes (1973) の *Guttman-Lingoes Program Series* としてまとめられている。例えば、SSA-II は、非対称行列に対して、行解と列解という 2 種類の解を求めることにより対応する。この点で、SSA-II は非対称 MDS の 1 つと考えられなくはない。

その後、後で述べる Takane, Young, & De Leeuw (1977) はトーガソン以降に提案された多くのモデルを統一的なアルゴリズムで解く ALSCAL を提案したが、その中には、クラスカルの単調回帰のアルゴリズムを用いて順序データの場合の対象の布置を求めるアルゴリズムも含まれている。一方、SAS では MDS プロシジャにより、最適化の方法が ALSCAL とは若干異なるが、ALSCAL の多くのモデルが適用できる。また、やほりのちに最尤 MDS の 1 つとして述べる Takane (1981) の方法でも、ノンメトリック MDS がその中の 1 つのオプションとして組み込まれている。

### 3.1.2 個人差 MDS

対象相互の非類似度行列が、個人により異なる場合、それらを  $\Delta_1, \Delta_2, \dots, \Delta_m$  とする。各行列は一般に  $n$  個の対象（刺激）に関する  $n$  行  $n$  列の非類似度行列である。そのような個人差を考慮する MDS は、個人差 MDS と呼ばれる。よく知られた個人差 MDS には、既に 1.3.2 節でその名前を紹介した Tucker & Messick (1963) の PVA と、Carroll & Chang (1970) の INDSCAL (INDividual Differences multidimensional SCALing) がある。

#### PVA

PVA は、まず複数の非類似度行列のそれぞれを  $n(n-1)/2$  行からなる列ベクトルとして配列し直し、それらを  $m$  個分並べて  $n(n-1)/2$  行  $m$  列の行列  $X$  を作成する。つぎに、内積行列  $\Sigma = X^t X$  を計算し、エッカート・ヤングの定理により分解し、 $m$  次以下の（ベストフィット）行列  $\hat{X}_p$  を得る。最後に行列  $\hat{X}_p$  を単純構造となるよう回転し  $F_p$  とする。得られた行列の各列は非類似度判断の1つの「視点」(a view point) と見なされる。PVA の第1ステージは、これらの「視点」を抽出することにある。第2ステージではある種の折衷行列  $T_a, a = 1, 2, \dots, p$  が  $F$  の  $(i, a)$  要素  $f_{ia}$  を用いて、つぎのように計算される。

$$T_a = \frac{1}{m} \sum_{i=1}^m f_{ia} \Delta_i, \quad (3.4)$$

最後に、得られた上記  $T_a, a = 1, 2, \dots, p$  のそれぞれに対して通常の MDS が施され、 $p$  個の「視点」に対応する  $p$  個の対象の布置が得られる。

Ross (1966) は、PVA は時々負のウエイト  $f_{ia}$  が含まれ、その結果として非ユークリッド解が得られることを指摘している。また、Meulman & Verbon (1993) は、この負のウエイトを避ける手続きを提案している。

#### INDSCAL

一方、Carroll と Chang は、通常の MDS における対象を埋め込む空間（これを彼らは刺激空間 stimulus space と呼ぶ）は個人を越えて共通であり、個人による対象相互の非類似度の差異は刺激空間に対するウエイトが個人によって異なることによるとするモデルを考えと呼んだ。INDSCAL では、データは原則的に比率尺度レベルの非類似度データとして得られていることを仮定し、各個人ごとの非類似度行列を (2.14) 式を用いて内積行列

$$B_i = \{b_{ijk}\}, \quad i = 1, \dots, m, \quad (3.5)$$

に変換し、次の量

$$\phi^2 = \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^n \left( b_{ijk} - \sum_{t=1}^r w_{it} x_{jt} x_{kt} \right)^2, \quad (3.6)$$

を最小にするような布置を求める。結局このモデルでは、重み付きユークリッド距離 (weighted Euclidean distance)

$$d_{ijk} = \sqrt{\sum_{t=1}^r w_{it} (x_{jt} - x_{kt})^2}, \quad (3.7)$$

を仮定していることになる。

このモデルから、まず対象の解布置は次元ごとの伸縮を重みで吸収できるので、単位の不定性はあることが見て取れる。これは、布置座標を次元ごと基準化すれば回避できる。

一方、座標系の直交回転に対しては、重みで吸収することは特別な場合（軸の置換の場合）を除き不可能なので、回転の不定性は特別な場合を除きないことがわかる。また、上記モデルから、各対象の位置座標



は平行移動に対して不定性を持つが、内積に対応するデータが直接得られている場合は、布置の基準化により一意的に布置や重みを決定できる。非類似度データが比率尺度データとして得られている場合も、内積を原点が重心となるように変換すれば、原点及び回転に対する一意的な解を得られる。詳細については、例えば斉藤 (1980, pp.143-144) を参照のこと。

最後に、INDSCAL における個人の各次元上の重み  $w_{it}$  には正なる条件は課されていないが、経験的にはこのモデルがデータに適合することが多いという。またもし負の重みが得られる場合には、データがモデルに適合しないと見なすか、それとも負の重みが少数個であればそれらをゼロとおいて計算を繰り返し正の領域で解を止めるのも一案とされる (例えば、斉藤, 1980, p.151)。

次節で述べる ALSICAL では、間隔尺度及び比率尺度等の場合に拡張された INDSCAL モデルがオプションの1つに組み込まれている。

## INDSCAL の適用例

図 2.3 Carroll & Chang (1970) の 12 カ国の刺激布置と被験者布置

### 3.1.3 ALSICAL

Takane et al. (1977) は、Torgerson 以来の各種データ及び名義データに対する広範なモデルを交互最小2乗法 (alternating least squares method) で解くアルゴリズムを提案した。彼らの方法は ALSICAL (Alternating Least Squares algorithm for individual differences sCALing) と呼ばれ、つぎの量を最小化する:

$$\tau^2 = \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^n (\hat{d}_{ijk}^2 - d_{ijk}^2)^2. \quad (3.8)$$

ここで、 $\hat{d}_{ijk}^2$  は、特定の観測特性の下で最適変換されたデータである。また、ここでの  $d_{ijk}$  は、(3.7) 式のそれに等しい。もちろん、重み付きユークリッドモデルのウエイトがすべて等しいときは、単純ユークリッドモデルに帰着される。

ALSICAL におけるモデルの適合度指標をクラスカルのストレス STRESS 指標と比べると、形式的にもかなり異なることが明らかである。形式的な違いの1つは、後者には分母があるのに対して前者にはない点である。さらに、前者では適合度指標は、距離  $d_{jk}$  の二乗に対する ディスパリティ (disparity) すなわち  $\hat{d}_{jk}$  の二乗の差の二乗和 と定義されるのに対して、後者では 距離とディスパリティの差の二乗和 が分子にきている。SSTRESS は、この違いを考慮して高根らが (3) 式、本文では (3.8) 式の指標に対してつけた名前である。

ALSICAL の目的は、彼らが原論文で述べているように、キャロルらの (計量的) 個人差 MDS に対する頑健で効率的なノンメトリック版を提案することであった。本文中の INDSCAL における適合度指標 (2.6) 式をキャロルはストレイン (STRAIN) と呼ぶが、高根らはこの指標はデータが計量的な場合には直接的だがノンメトリックなデータの場合には大変複雑になる点を考慮して、(3.8) 式の SSTRESS 基準を提案した。

ALSICAL では、上記  $\tau^2$  を  $\hat{d}_{ijk}^2$  と  $d_{ijk}^2$  に関して交互に最小化しながら、全体として最小を求める方式を取る。ALSICAL では、データのタイプには比率尺度、間隔尺度、順序尺度、名義尺度の4つのオプションが、データの形には対称で非類似度、対称で類似度、非対称で非類似度、非対称で類似度、矩形で非類似度、矩形で類似度の6つのオプションが、測定条件にはマトリックス条件付き、行条件付き、無条件のデータ行列についての3つのオプション等が、選べる。また、モデルタイプとして、単純ユークリッドモデル、個人差モデル、多重過程非対称モデル (multiprocess asymmetric model)、多重過程非対称個人差モデル (multiprocess asymmetric individual differences model) が選べる (ALSICAL・4 ユーザズガイド (Young

& Lewyckyj, 1979)。ここで、多重過程とは非類似度行列の1つの生成過程に対してつけられた概念であり、(評定者としての)各対象に対応する(例えば)各行の非類似度が異なる被験者により生成された場合を指す。これに対して、すべての行が同一被験者により生成された場合は、単一過程 (single process) と呼ばれる (Young, 1975)。

一方、SASのMDSプロシジャでは、ALSCAL及び次節で述べるラムゼイのMULTISCALE (Ramsay, 1977, 1982)の多くの特徴を実現しているが、未知パラメータの推定方法に関しては、非線形最小2乗法を用いており、ALSCALの交互最小2乗法と若干異なる (SAS Institute, 1999)。データのタイプ (LEVEL オプション) は、絶対尺度、比率尺度、間隔尺度、対数間隔尺度、順序尺度の5つのオプションが用意されている。一方、データの型 (SHAPE オプション) では三角型、正方形の2つのオプションがある。また、行列のタイプ (COEF オプション) として単位、及び対角型があり、それぞれ単純ユークリッド距離モデルや重み付きユークリッド距離モデルに対応している。Version 8のオンラインドキュメントによれば、MDSプロシジャでは、とりわけFITオプションで2乗距離 (SQUARED) を選択し、さらに1、2のオプションを選択すると、ALSCALとほぼ同様の結果が得られるという (SAS Institute, 1999)。

### 3.1.4 最尤 MDS

Takane et al. (1977)までの従来のMDSでは、未知数としての対象の多次元空間での座標値の推定には、最小2乗法もしくは固有値問題の解法が用いられてきた。ここで、固有値問題の解も最小2乗的特性を持つ (Stewart & Sun, 1990) ことに注意すると、従来のMDSでの座標値は、最小2乗的に推定されてきたと言ってよい。

これに対して、70年代の後半からは、座標値の推定に最尤法 (maximum likelihood method) を用いるMDSが幾つか提案されだした (Ramsay, 1977, 1978, 1982; Takane, 1978a, 1978b, 1981; Takane & Carroll, 1981)。ここでは、まずその中で礎石となった Ramsay (1977) について紹介する。

最尤法を用いるMDS、すなわち最尤MDS (maximum likelihood MDS) では、方法の性格上誤差に対して何らかの分布形を仮定する必要がある。この問題については、もちろん Ramsay (1977) 以前にも、たとえば対象の多次元空間での座標値に正規分布を仮定したモデルが提唱されている (Suppes & Zinnes, 1963; Ramsay, 1969)。しかしながら、このアプローチは、Zinnes and Griggs (1974) が、彼らの確率的展開法 (probabilistic unfolding) を定式化するまで、まったく成功の例を見ていなかった。

## MULTISCALE

これに対して Ramsay (1977) は、観測非類似度  $\delta_{jkr}$  に対して対数正規分布の仮定を置き、最尤法により対象の座標値を推定する方法を提案した。すなわち、この方法ではまず  $\delta_{jkr}$  の分布を次式で定義する：

$$\ln \delta_{jkr} \sim N(\ln d_{jk}, \sigma^2). \quad (3.9)$$

ここで、添字  $r$  は、対象の対  $(j, k)$  の非類似度に対する第  $r$  番目の繰り返し判断であることを表す。

この時、 $\delta_{jkr}$  の密度関数は

$$f(\delta_{jkr}) = \frac{1}{\sqrt{2\pi} \sigma \delta_{jkr}} \exp \left\{ -\frac{[\ln(\delta_{jkr}/d_{jk})]^2}{2\sigma^2} \right\}, \quad (3.10)$$

で与えられる。

これより、対数正規分布の仮定のもとでデータの得られる尤度は

$$L = \prod_{j=1}^n \prod_{k=1}^n \prod_{r=1}^R f(\delta_{jkr}), \quad (3.11)$$

で与えられる。もちろん尤度の定義から、 $L$  は欠測値を除いたものとして差し支えない。

Ramsay は、彼のプログラムを MULTISCALE と呼ぶ。MULTISCALE における距離の定義は、INDSCAL や ALSCAL と同じ (3.7) 式の重み付きユークリッド距離に等しく、ALSCAL と同様その特殊ケースとして単純ユークリッド距離モデルを含む。既述の (3.7) 式は、右辺を行列表現すれば、

$$d_{ijk} = \sqrt{(\mathbf{x}_j - \mathbf{x}_k)^t \mathbf{W}_i (\mathbf{x}_j - \mathbf{x}_k)}, \quad (3.12)$$

と書けるが、Ramsay (1982) は、これを行列  $\mathbf{W}_i$  が単位行列、対角行列 (diagonal matrix)、もしくは半定符号かにより、単位計量 (identity metric)、対角計量 (diagonal metric)、フルメトリック (full metric) の3種類に分類して区別している。

最尤 MDS の最小 2 乗 MDS に比べてのメリットは、次元数の統計的検定や各対象の多次元空間内での推定座標値に対する信頼領域 (confidence region) を描ける点であろう。もちろん、これらのメリットを享受するためには、データが分布の仮定に適合していなければならない。最尤 MDS における対象の布置座標値の信頼領域の推定の基礎は、一般に母数  $\boldsymbol{\theta}_t = (\theta_1, \theta_2, \dots, \theta_k)^t$  に対する (同時) 最尤推定量  $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k)^t$  が、漸近的に平均  $\boldsymbol{\theta}_t$ 、共分散行列  $\mathbf{I}^{-1}(\boldsymbol{\theta}_t)$  を持つ多変量正規分布に従うことを利用することにある。ここで、行列  $\mathbf{I}(\boldsymbol{\theta}_t)$  は、情報行列 (information matrix) であり、

$$\mathbf{I}(\boldsymbol{\theta}_t) = \left\{ E \left( \frac{\partial \ln L(\boldsymbol{\theta})}{\partial \theta_r} \frac{\partial \ln L(\boldsymbol{\theta})}{\partial \theta_s} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_t} \right\} = \left\{ -E \left( \frac{\partial^2 \ln L}{\partial \theta_r \partial \theta_s} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_t} \right\} \quad (3.13)$$

実際には、この種の MDS では未知パラメータである対象の布置座標は座標軸の回転と移動に関して自由なので、この結果をそのまま (漸近的) 信頼領域の計算に用いることはできず、パラメータの推定時に工夫が必要である。詳細は Ramsey (1978) を参照のこと。

## MAXSCAL

一方、Takane (1981) はトーガソンのカテゴリー判断の法則を利用し、系列カテゴリー法 ( $M$  個のカテゴリーから成るとして) による対象相互の (非) 類似度判断データから、最尤法により対象の多次元布置を求める方法 MAXSCAL (MAXSCAL2.1) を提案した。ここで、対象対  $(O_j, O_k)$  に対する被験者  $i$  の機会  $t$  での類似度 (観測 (非) 類似度そのものではない) を  $\lambda_{jki}^{(t)}$  と書くとして、高根は表現モデルとして単純ユークリッド距離モデルを仮定し、さらに加法的モデル (additive model)

$$\lambda_{jki}^{(t)} = d_{jk} + \varepsilon_{jki}^{(t)}, \quad \varepsilon_{jki}^{(t)} \sim N(0, \sigma_k^2), \quad (3.14)$$

と乗法的モデル (multiplicative model)

$$\lambda_{jki}^{(t)} = d_{jk} \varepsilon_{jki}^{(t)}, \quad \ln \varepsilon_{jki}^{(t)} \sim N(0, \sigma_k^2), \quad (3.15)$$

を考える。後者の誤差モデルは、ラムジーの最尤 MDS における対数正規分布を仮定したモデルと同じである。

つぎに、系列カテゴリー間の境界値を小さい順に  $b_{i1}, b_{i2}, \dots, b_{iM-1}$  とし、さらに  $b_{i0} = -\infty, b_{iM} = \infty$  と書けば、被験者  $i$  が対象対  $(O_j, O_k)$  の (非) 類似度が第  $m$  カテゴリーに属すると判断される確率  $P_{jkim}$  は、

$$P_{jkim} = \int_{b_{i,m-1}}^{b_{im}} g(\lambda_{jki}) d\lambda_{jki}. \quad (3.16)$$

ここで、 $g(\lambda_{jki})$  は  $\lambda_{jki}$  の密度関数とする。この時、例えば加法的モデルの場合、

$$P_{jkim} = \int_{a_{jki,m-1}}^{a_{jki,m}} f(z) dz. \quad (3.17)$$

ここで、 $f(z)$  は単位正規分布の密度関数であり、

$$z = (\lambda_{jki} - d_{jk})/\sigma_i, \quad a_{jki,m} = (b_{im} - d_{jk})/\sigma_i. \quad (3.18)$$

(3.17) 式の  $P_{jkim}$  は  $f(z)$  の分布関数を  $F_{jkim}$  と書けば、 $P_{jkim} = F_{jkim} - F_{jki\ m-1}$  と書けることに注意しよう。

つぎに、 $\delta_{jkir}$  を対象対  $(O_j, O_k)$  に対する被験者  $i$  の第  $r$  番目の観測 (非) 類似度判断として、つぎの  $Z_{jkimr}$  を導入しよう：

$$Z_{jkimr} = \begin{cases} 1 & (\delta_{jkir} \text{ がカテゴリ } m \text{ に属する時}) \\ 0 & (\text{それ以外の時}) \end{cases} \quad (3.19)$$

この時、対象対に対する特定の評定判断の確率は、 $\prod_{m=1}^M P_{jkim}^{Z_{jkimr}}$  と書けるので、全観測値の同時尤度は

$$L = \prod_i \prod_{j,k} \prod_r \prod_m P_{jkim}^{Z_{jkimr}} = \prod_i \prod_{j,k} \prod_m P_{jkim}^{Y_{jkim}}. \quad (3.20)$$

ここで、 $Y_{jkim} = \sum_r Z_{jkimr}$  とする。この  $Y_{jkim}$  は対象対  $(O_j, O_k)$  に対する (非) 類似度判断が、カテゴリ  $m$  に落ちる度数である。

これより、対象布置座標  $\{x_{ja}\}$ 、ばらつき  $\sigma_i$ 、カテゴリの境界値  $\{b_{im}\}$  の最尤解は、

$$\ln L = \sum_i \sum_{j,k} \sum_m Y_{jkim} \ln P_{jkim}, \quad (3.21)$$

を最大にするようなパラメータとして求められる。

未知パラメータ  $b_{im}$  に対して、いろいろな制約を加えることにより、われわれは 順序尺度データ や 間隔尺度データ 等に対する最尤解を求めることができる。高根は、MAXSCAL2.1 における尤度比検定や AIC を用いた複数のモデルの優劣の検定や比較について述べている。詳細は、Takane (1981) を参照のこと。

Chino (1992) は、MAXSCAL を非対称 MDS の場合に拡張する試みを提案している。また、Saburi & Chino (2004) は、これを発展させ、原理的には任意の非対称 MDS モデルについて、順序尺度レベルから比率尺度レベルまでの尺度レベルのどの尺度レベルの非対称関係行列に対しても、対称性の検定 (test for symmetry)、各対象の信頼領域の描画、AIC によるモデル間比較等を可能にするアルゴリズムを提案する。

### 最尤 MDS の適用例

図 2.4 Ramsey (1978) の最尤 MDS による刺激座標の漸近的信頼領域の例

## 3.2 その他の MDS と最近の展開

### 3.2.1 多次元展開法

多次元尺度構成法では、データは通常複数の対象間の (非) 類似度 (判断) を正方行列の形にまとめたものであるが、これとは形式的には異なる、複数刺激  $S_1, S_2, \dots, S_n$  に対する個人  $O_1, O_2, \dots, O_m$  の選好 (preference) データから、刺激及び個人を同一の多次元空間上に位置づける方法もあり、(多次元) 展開法と呼ばれる。

展開法は、Coombs (1950, 1964) により始められたもので、最初は刺激に対する個人の選好順位のデータから、1次元順序尺度を構成する方法として出発し、ベネットとヘイズにより多次元順序尺度構成の方法に発展し (Benett & Hays, 1960; Hays & Bennett, 1961)、現在では古典的展開法と呼ばれる (斉藤, 1980)。古典的展開法は、順序尺度の構成法であり、厳密にはここで議論している (距離空間にデータをフィットさせる) 狭義の多次元尺度構成法の範疇には入らない。

しかしながら、古典的展開法はその後、選好順位データをもとに刺激布置及び理想点の多次元距離空間上の布置を仮定し、順位データが可能な限り点間距離の単調減少関数となるように両布置を定める 非計量

的展開法 (nonmetric unfolding) へと発展した (Lingoes, 1966; Kruskal, 1968; Young & Torgerson, 1967)。また、シェネマンらにより、計量的展開法 (metric unfolding) として発展しており (Greenacre & Browne, 1986; Schönemann, 1972; Schönemann & Wang, 1972; Zinnes & Griggs, 1974)、多次元尺度構成法の1つと見なせよう。また、計量的データの場合にも非計量的データの場合にもデータを距離空間上に位置づける統一的なアルゴリズムを提案した Takane, Young, & de Leeuw (1977) による ALSCAL でも1つのモデルとして組み込まれている。

### 古典的展開法

そこで、まずここでは古典的展開法とりわけクームズの1次元展開法について簡単にふれる。彼の展開法では、刺激と個人はJ尺度と呼ばれる共通の1次元尺度上の点で表され、それぞれの個人の刺激に対する選好順位は各個人の理想点 (the ideal point) と呼ばれる位置から刺激点までの距離の順位に一致する、と仮定される。

一方、個人の刺激に対する選好順位はI尺度と呼ばれ、図2.5のようにJ尺度上の個々の刺激を個人  $O_k$  の理想点  $IP_k$  で折りたたむ (fold) ことによって得られるとされる。

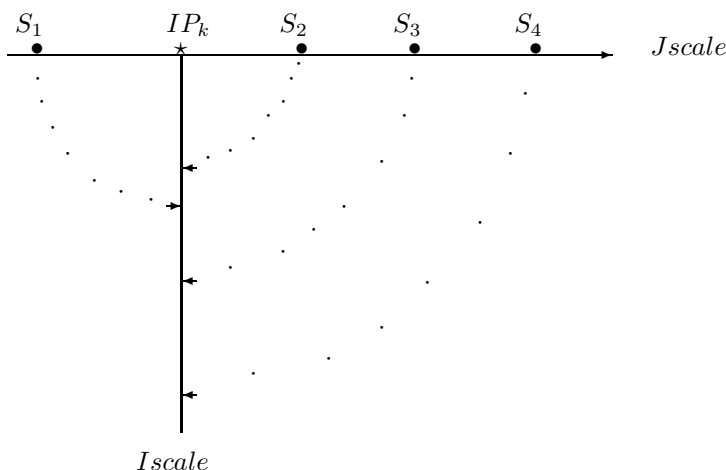


図 3.5: クームズの展開法における J 尺度と I 尺度の関係と理想点  $IP_i$

実際の作業は、刺激に対する (I 尺度上の) 選好順位データから J 尺度 (上での刺激の位置と個人の理想点) を求めることになる。クームズの展開法はその後、ベネットとヘイズ (Bennett & Hays, 1960; Hays & Bennett, 1961) により、多次元順序尺度の場合に拡張された。これらの方法は、現在では古典的展開法と呼ばれている (斉藤, 1980)。

古典的展開法は、その後選好順位データをもとに、刺激布置と理想点の多次元距離空間上の布置座標を仮定し順位データが可能な限り点間距離の単調減少関数になるように両布置を定める非計量的展開法 (nonmetric unfolding) へとさらに発展している (Lingoes, 1966; Kruskal, 1968; Young & Torgerson, 1967)。

### 計量的展開法

一方、シェネマン (Schönemann, 1972) は、刺激と個人の理想点間の距離の推定値が比率尺度で得られている場合の計量的展開法 (metric unfolding) を提案した。計量的展開法では、刺激布置  $X = \{x_{ia}\}$  と個

人の理想点布置  $Y = \{y_{ka}\}$  間の距離  $d_{ik}$  の推定値が何らかの方法で得られている場合に、刺激布置  $X$  及び理想点の布置  $Y$  を求める方法である。

まず、

$$d_{ik}^2 = \sum_{a=1}^p (x_{ia} - y_{ka})^2 = \sum_{a=1}^p x_{ia}^2 - 2 \sum_{a=1}^p x_{ia}y_{ka} + \sum_{a=1}^p y_{ka}^2, \quad i = 1, 2, \dots, n, k = 1, 2, \dots, m, \quad (3.22)$$

と書き、行列表現すると

$$D^{(2)} = \{d_{ik}^2\} = \text{diag}(XX^t)\mathbf{1}_n\mathbf{1}_m^t - 2XY^t + \mathbf{1}_n\mathbf{1}_m^t \text{diag}(YY^t), \quad (3.23)$$

と書ける。ここで、 $\mathbf{1}_n$  及び  $\mathbf{1}_m$  は 1 をそれぞれ  $n$  個及び  $m$  個並べた列ベクトルとする。

計量的展開法では、まずこの  $D^{(2)}$  を 2 種類の中心化行列を用いて

$$P = -\frac{1}{2} G_n D^{(2)} G_m = G_n X Y^t G_m, \quad (3.24)$$

と変換し、古典的 MDS の場合と同様、 $P = X^*(Y^*)^t$  と分解する。実際には、エッカート・ヤングの方法で  $P$  の  $p$  次元のベストフィットを行い、 $P = U V^t$  とする。この時、 $X^*$  及び  $Y^*$  は適当な正則行列  $T$  を用いて

$$X^* = UT, \quad Y^* = V(T^{-1})^t, \quad (3.25)$$

と書ける。

ここで、 $X^*$  と  $Y^*$  はそれぞれ異なる原点を持つ布置なので、共通の原点を持つ布置に変換する必要がある。このために、 $y_0$  を共通の原点を刺激布置の重心に移動するために必要な移動量を表すベクトルとすると、

$$X = X^* = UT, \quad Y = Y^* + \mathbf{1}_m y_0^t = V(T^{-1})^t + \mathbf{1}_m y_0^t, \quad (3.26)$$

と書ける。データから上記  $T$  と  $y_0$  を推定できれば、われわれは理想点布置  $X$  と刺激布置  $Y$  を推定できることになる。シェーネマンはそのための 1 つのアルゴリズムを示した。しかし、Gold (1973) は、シェーネマンのアルゴリズムが布置の変換に必要な方程式として非線形部を無視している点や、一般的にはシェーネマンの計量的展開法は大局的な解を必ずしも持たない点などを指摘した。

のちに述べる高根らの ALSCAL では、計量的展開法も非計量的展開法も彼らのアルゴリズムを用いて実現しており、SAS の MDS プロシジャでも利用できる。この高根らのアルゴリズムを改善する試みが、その後 Greenacre & Browne (1986) により提案された。また、展開法は、選好順位データに誤差分布を仮定する確率的展開法へも発展している (Schönemann & Wang, 1972; Zinnes & Griggs, 1974)。

### 3.2.2 内積モデル

距離モデルに対するもう 1 つのモデルに内積モデル (inner product model) がある。このモデルを最初に提案したのは、Ekman (1954) である。彼は、14 色の周波数の異なる色の対を簡単な 5 段階評定尺度で被験者に評定させて得た類似度得点を 0 から 1 点の値に変換し、グループセントロイド法で因子分析した。その結果、因子負荷量と周波数との間に興味深い関係のあることを見つけた。

もっとも、彼の方法では類似度行列の対角要素は観測されていないこと、及び尺度が単純な 5 段階評定尺度であることから、彼が対象を埋め込む空間として距離空間を考えていた可能性は低いと言わねばなるまい。この意味では、この時点での彼の方法は、前節で定義した狭義の意味では MDS とは言えないことになる。

エックマンはその後、彼の方法をスティーブンスの 1 次元比率尺度構成法の多次元への拡張、すなわち多次元比 (率) 尺度構成法 (multidimensional ratio scaling) と位置づけ、以下のような対象間の内積行列  $R = \{r_{jk}\}$  を計算する方法を示した。すなわち、 $n$  個の刺激から 2 対づつ刺激を選び、 $O_j$ 、 $O_k$  の原

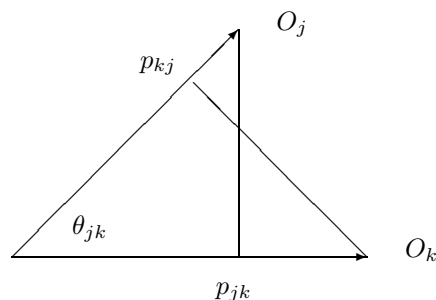


図 3.7:  $O_j$ 、 $O_k$  の、他方から見た時の成分  $p_{jk}$ 、 $p_{kj}$

点からの距離を  $h_j$ 、 $h_k$  とし、それぞれの刺激を他方から見た時の成分  $p_{jk}$ 、 $p_{kj}$  を比推定させ、それらを  $q_{jk} = p_{jk}/h_k$ 、 $q_{kj} = p_{kj}/h_j$  となったとする。図 2.7 から  $p_{jk} = h_j \cos \theta_{jk}$  なので、

$$q_{jk} = \frac{h_j}{h_k} \cos \theta_{jk}. \quad (3.27)$$

これより、2つのベクトルの内積  $r_{jk}$  は、

$$r_{jk} = q_{jk} h_k^2 = h_j h_k \cos \theta_{jk}. \quad (3.28)$$

さらに、(3.27) 式より、

$$\frac{q_{jk}}{q_{kj}} = \frac{h_j^2}{h_k^2}. \quad (3.29)$$

そこで、 $\sum_{k=1}^n (1/h_k^2) = 1$  と測定の単位を取れば、

$$h_j^2 = \sum_{k=1}^n \frac{q_{jk}}{q_{kj}}. \quad (3.30)$$

Micko & Lehmann (1969) は、 $q_{jk}$  が誤差ありデータの場合、 $r_{jk}$  や  $h_k^2$  を推定するための2種類の最小2乗法を提案している。Micko (1970) は、さらにエックマンにおける  $p_{jk}$  が正で  $h_j$  以下なる制約条件をはずした場合の多次元比(率)尺度構成法を提案した。

### 3.2.3 非対称 MDS

多次元展開法等の特別な例を除き、通常 MDS では対象間に観測される(非)類似度データ行列は対称である。すなわち、任意の対象  $O_j$  から  $O_k$  への(非)類似度と  $O_k$  から  $O_j$  への(非)類似度は等しいとの仮定がなされている。したがって、非対称なデータ行列についての MDS、すなわち非対称 MDS (asymmetric MDS) が開発されるまでは、対象間で観測される(非)類似度の非対称性は誤差と見なされ両観測値の平均を取ればよいとされてきた。

しかし、現実には対象相互の(非)類似度が誤差以上のものである場合は数多く存在する。そのような場合の MDS、非対称 MDS は Young (1975) 以来現在までに数多く提案されてきている(例えば、千野、

1997; Cox & Cox, 2001)。非対称 MDS では、ユークリッド空間の修正版や佐藤 (1989) による非対称ミンコフスキー空間 (asymmetric Minkowski space)、Chino & Shiraiwa (1993) による有限次元 (複素) ヒルベルト空間 (finite-dimensional (complex) Hilbert space) などが用いられる。紙面の都合上、非対称 MDS の詳細は、千野 (1997) を参照のこと。

Chino & Shiraiwa (1993) は、比率尺度レベルで得られた非対称関係データ行列をエルミート行列に変換すれば、エルミート行列が正定符号の時は対象を有限次元複素ヒルベルト空間に、不定符号の時は対象を不定計量空間に位置づけることができることを示した。その中核に位置付けられる千野・白岩の定理は、ヤング・ハウスホルダーの定理を複素空間の場合に拡張するものである。

定理 2 (Chino-Shiraiwa)  $d_{jk} = \|\mathbf{v}_j - \mathbf{v}_k\|$  および  $\bar{d}_{jk} = \|\mathbf{v}_j - i\mathbf{v}_k\|$  が (有限次元) ヒルベルト空間上の真の点の集合 (組) 間の相互距離であるための必要十分条件は、エルミート内積から成るつぎのエルミート行列  $\mathbf{H}$  が正の半定符号であることである。

$$\mathbf{H} = \left\{ \frac{1}{2}(d_{j0}^2 + d_{k0}^2 - d_{jk}^2) + \frac{1}{2}i(d_{j0}^2 + d_{k0}^2 - \bar{d}_{jk}^2) \right\}, \quad (3.31)$$

[演習 2.1] 千野・白岩の定理を証明せよ。

小杉らは、社会心理学におけるバランス理論を、千野・白岩の定理に基づくモデル HFM (Hermitian Form Model) を適用することにより、非対称な場合に拡張している (小杉他, 2004a; 小杉, 2004b)。

非対称 MDS は、ヤング以来このように多くのモデルが提案されてきているが、ノンメトリック非対称 MDS については未だ Okada & Imaizumi (1987, 1997) ぐらいしか提案されていないし、非対称最尤 MDS も雑誌に公刊されたものは特別なモデルを除き (Adachi, 1999, 2003; de Rooij & Heiser, 2003)、ほとんどないのが現状である (Chino, 2003)。近々、Saburi & Chino (2004) は、対称性の検定も含む、広範な非対称 MDS モデルに対する高根の MAXSCAL を拡張する非対称最尤 MDS を提案する。

これまでに提案されてきた非対称 MDS も含めた MDS では、対象の布置は観測 (非) 類似度データ行列に対する最小二乗解や単調回帰解、あるいは最尤解など、あくまでも統計的な意味でのデータの整合性を基準として推定される。しかしそのようなモデルには、対象相互の観測 (非) 類似度についての心理学的知見が盛り込まれることはほとんどないと言ってよい。また、例えばある時点までの対象相互の観測 (非) 類似度データ行列からつぎの時点の (非) 類似度を、心理学的な知見をモデルに組み込んで予測するなどのモデルもない。最近、千野はそのようなモデルを提案しつつある (Chino, 2002, 2003)。



## References

- Adachi, K. (1999). Constrained multidimensional unfolding of confusion matrices: Goal point and slide vector models. *Japanese Psychological Research*, **41**, 152-162.
- Adachi, K. (2003). A latent variable model for multidimensional unfolding. In H. Yanai, A. Okada, K. Shigemasu, Y. Kano, & J. J. Meulman (Eds.) *New Developments in Psychometrics*, Tokyo: Springer. pp.503-510.
- Arbuckle, J., & Nugent, J. H. (1973). A general procedure for parameter estimation for the law of comparative judgement. *British Journal of Mathematical and Statistical Psychology*, **26**, 240-260.
- Bennett, J. F., & Hays, W. L. (1960). Multidimensional unfolding: Determining the dimensionality of ranked preference data. *Psychometrika*, **25**, 27-43.
- Blank, A. A. (1953). The Luneburg theory of binocular visual space. *Journal of the Optical Society of America*, **43**, 717-727.
- Bock, R. D. and Jones, L. V. (1968). *The Measurement and Prediction of Judgment and Choice*. San Francisco: Holden-Day.
- Bogardus, E. S. (1925). Measuring social distance. *Journal of Applied Sociology*, **9**, 299-308.
- Borg, I., and Shye, S. (1995). *Facet Theory - Form and Content*, London: Sage Publications.
- Carroll, D. J., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an  $N$ -way generalization of "Eckart-Young" decomposition. *Psychometrika*, **35**, 283-319.
- Carnap, R. (1966). *Philosophical Foundations of Physics*. New York: Basic Books.
- 千野直仁 (1997). 非対称多次元尺度構成法 現代数学社
- Chino, N. (2002). Complex space models for the analysis of asymmetry. In S. Nishisato, Y. Baba, H. Bozdogan, & K. Kanefuji (Eds.) *Measurement and Multivariate Analysis*, Tokyo: Springer. pp.107-114.
- Chino, N. (2003). Complex difference system models for the analysis of asymmetry. In H. Yanai, A. Okada, K. Shigemasu, Y. Kano, & J. J. Meulman (Eds.) *New Developments in Psychometrics*, Tokyo: Springer. pp.479-486.
- Chino, N., & Shiraiwa, K. (1993). Geometrical structures of some non-distance models for asymmetric MDS. *Behaviormetrika*, **20**, 35-47.
- Coombs, C. H. (1950). Psychological scaling without a unit of measurement. *Psychological Review*, **57**, 145-158.
- Coombs, C. H. (1964). *Theory of Data*. New York: Wiley.
- Cox, T. F. and Cox, M. A. A. (2001). *Multidimensional Scaling*. 2nd Edition. London: Chapman & Hall.
- De Rooij, M., & Heiser, W. J. (2003). A distance representation of the quasi-symmetry model and related distance models. In H. Yanai, A. Okada, K. Shigemasu, Y. Kano, & J. J. Meulman (Eds.) *New Developments in Psychometrics*, Tokyo: Springer. pp.487-494.
- Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, **1**, 211-218.
- Ekman, G. (1954). Dimensions of color vision. *The Journal of Psychology*, **38**, 467-474.
- Ekman, G. (1963). A direct method for multidimensional ratio scaling. *Psychometrika*, **28**, 33-41.
- Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf and Haertel.
- Gallant, A. R. (1987). *Nonlinear Statistical Models*. New York: Wiley.
- Gold, E. M. (1973). Metric unfolding: Data requirement for unique solution & clarification of Schönemann's algorithm. *Psychometrika*, 555-569.

- Greenacre, M. J., & Browne, M. W. (1986). An efficient alternating least-squares algorithm to perform multidimensional unfolding. *Psychometrika*, **51**, 241-250.
- Guilford, J. P. and Fruchter, B. (1973). *Fundamental Statistics in Psychology and Education*. London: McGraw-Hill.
- Guttman, L. (1944). A basis for scaling quantitative data. *American Sociological Review*, **9**, 139-150.
- Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika*, **33**, 469-506.
- Hays, W. L., & Bennett, J. F. (1961). Multidimensional unfolding: Determining configuration from complete rank order preference data. *Psychometrika*, **26**, 221-238.
- Hohle, R. H. (1966). An empirical evaluation and comparison of two models for discriminability scales. *Journal of Mathematical Psychology*, **3**, 174-183.
- Horst, P. (1965). *Factor Analysis of Data Matrices*, New York: Holt, Rinehart and Winston, Inc.
- Indow, T. (1983). Multidimensional mapping in Riemannian space and its application. *Paper presented at the annual meeting of the Psychometric Society*, Los Angeles, CA.
- Indow, T. (1988). Multidimensional studies of Munsell color solid. *Psychological Review*, **95**, 456-470.
- Indow, T., & Ohsumi, K. (1972). Multidimensional mapping of sixty Munsell colors by non-metric procedure. In J. J. Vos, L.F. C. Friele, & P. L. Walraven (Eds.), *Color metrics: Proceedings of the Helmholtz Memorial Symposium on Color Metrics* (pp. 124-133). the Netherlands: Institute for Perception.
- Indow, T., & Shiose, T. (1956). An application of the method of multidimensional scaling to perception of similarity or difference in colors. *Japanese Psychological Research*, **1**, 45-64.
- Indow, T., & Watanabe, T. (1988). Alleys on an extensive apparent frontoparallel plane: a second experiment. *Perception*, **17**, 647-666.
- 木村通治・真鍋一史・安永幸子・横田賀英子 (2002). ファセット理論と解析事例 ナカニシヤ
- 小杉考司・藤沢隆史・藤原武弘 (2004a). バランス理論と固有値分解 理論と方法 **35** (印刷中)
- 小杉考司 (2004b). 心理-論理と態度理論への数理アプローチ 関西学院大学大学院社会学研究科提出博士論文
- Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, **29**, 1-27.
- Kruskal, J. B. (1964b). Multidimensional scaling: A numerical method. *Psychometrika*, **29**, 115-129.
- Kruskal, J. B. (1968). How to use MDSCAL, a program to do multidimensional scaling and multidimensional unfolding. Bell Telephone Laboratories.
- Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. California: Sage Publications.
- Lancaster, P., & Tismenetsky, M. (1985). *The Theory of Matrices*, 2nd Edition, New York: Academic Press.
- Lawson, C. L. and Hanson, R. J. (1974). *Solving Least Squares Problems*. London: Prentice-Hall.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, No. **140**, 5-55
- Lingoes, J. C. (1966). An IBM program for Guttman-Lingoes smallest space analysis-R II. *Behavioral Science*, **11**, 322.
- Lingoes, J. C. (1973). *The Guttman-Lingoes Nonmetric Program Series*. Ann Arbor: Mathesis Press.
- Lunenburg, R. K. (1950). The metric of binocular visual space. *Journal of the Optical Society of America*, **40**, 627-643.
- Manabe, K. (2001). *Facet theory and studies of Japanese Society : from a comparative perspective*. Bonn: Bier'sche Verlagsanstalt.
- McCullagh, P. and Nelder FRS, J. A. (1989). *Generalized Linear Models*. London: Chapman and Hall.

- Messick, S. J., & Abelson, R. P. (1956). The additive constant problem in multidimensional scaling. *Psychometrika*, **21**, 1-15.
- Meulman, J. J., & Verboon, P. (1993). Points of view analysis revisited: Fitting multidimensional structures to optimal distance components with cluster restrictions on the variables. *Psychometrika*, **58**, 7-35.
- Micko, H. C. (1970). A "Halo"-Model for multidimensional ratio scaling. *Psychometrika*, **35**, 199-227.
- Micko, H. C., & Lehmann, G. (1969). Two least-squares solutions for Ekman's method of direct multidimensional ratio scaling. *Scandinavian Journal of Psychology*, **10**, 57-60.
- Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. New York: Wiley.
- 大山正・今井省吾・和気典二編 (1994). 新編 感覚・知覚心理学ハンドブック 誠信書房
- Okada, A., & Imaizumi, T. (1987). Nonmetric multidimensional scaling of asymmetric proximities. *Behaviormetrika*, **21**, 81-96.
- Okada, A., & Imaizumi, T. (1997). Asymmetric multidimensional scaling of two-mode three-way proximities. *Journal of Classification*, **14**, 195-224.
- Ramsay, J. O. (1969). Some statistical considerations in multidimensional scaling. *Psychometrika*, **34**, 167-182.
- Ramsay, J. O. (1977). Maximum likelihood estimation in multidimensional scaling. *Psychometrika*, **42**, 241-266.
- Ramsay, J. O. (1978). Confidence regions for multidimensional scaling analysis. *Psychometrika*, **43**, 145-160.
- Ramsay, J. O. (1982). Some statistical approaches to multidimensional scaling data. *The Journal of the Royal Statistical Society, Series A (General)*, **145**, 285-312.
- Richardson, M. W. (1938). Multidimensional Psychophysics. *Psychological Bulletin*, **35**, 659. (Abstract).
- Robertson, T., Wright, F. T., and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*. New York: Wiley.
- Ross, J. (1966). A remark on Tucker and Messick's "point of view" analysis. *Psychometrika*, **31**, 27-31.
- Saffir, M. A. (1937). A comparative study of scales constructed by three psychophysical methods. *Psychometrika*, **2**, 179-198.
- Saburi, S., & Chino, N. (2004). A maximum likelihood for asymmetric MDS. *Paper to be presented at the 31th annual meeting of the Behaviormetric Society of Japan*, Yokohama, Japan. Saito, T. (1978). The problem of the additive constant and eigenvalues in metric multidimensional scaling. *Psychometrika*, **43**, 193-201.
- 齋藤堯幸 (1980). 多次元尺度構成法 朝倉書店
- SAS Institute (1999). *SAS/STAT User's Guide*, Version 8.1 Cary: SAS Institute Inc.
- 佐藤義治 (1989). Minkowski 計量を用いた非対称類似性の距離表現について 計算機統計学 第2巻, 35-45.
- Schönemann, P. H., & Tucker, L. R. (1967). A maximum likelihood solution for the method of successive intervals allowing for unequal stimulus dispersions. *Psychometrika*, **32**, 403-417.
- Schönemann, P. H., & Wang, M. M. (1972). An individual difference model for the multidimensional analysis of preference data. *Psychometrika*, **37**, 275-309.
- Shepard, R. N. (1962a). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, **27**, 125-140.
- Shepard, R. N. (1962b). The analysis of proximities: Multidimensional scaling with an unknown distance function. II. *Psychometrika*, **27**, 219-246.
- Stevens, S. S. (1951). Mathematics, measurement and psychophysics. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology*. New York: Wiley.
- Stewart, G. W., & Sun, J. (1990). *Matrix Perturbation Theory*. New York: Academic Press.

- Suppes, P., & Zinnes, J. L. (1963). Basic measurement theory. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of mathematical Psychology*, **1**, New York: John Wiley.
- Takane, Y. (1978a). A maximum likelihood method for nonmetric multidimensional scaling: 1. The case in which all empirical pairwise orderings are independent- Theory. *Japanese Psychological Research*, **20**, 7-17.
- Takane, Y. (1978b). A maximum likelihood method for nonmetric multidimensional scaling: 1. The case in which all empirical pairwise orderings are independent- Evaluation. *Japanese Psychological Research*, **20**, 105-114.
- 高根芳雄 (1980). 多次元尺度法 東京大学出版会
- Takane, Y. (1981). Multidimensional successive categories scaling: A maximum likelihood method. *Psychometrika*, **46**, 9-28.
- Takane, Y., & Carroll, D. J. (1981). Nonmetric maximum likelihood multidimensional scaling from directional rankings of similarities. *Psychometrika*, **46**, 389-405.
- Takane, Y., Young, F. W., & de Leeuw, J. (1977). Nonmetric individual differences multidimensional scaling: An alternative least squares method with optimal scaling features. *Psychometrika*, **42**, 1-68.
- 竹内啓編集 (1989). 統計学辞典 東洋経済
- 竹内啓・柳井晴夫 (1972). 多変量解析の基礎 東洋経済新報社
- Thurstone, L. L. (1925). A method of scaling psychological and educational tests. *Journal of Educational Psychology*, **16**, 433-451.
- Thurstone, L. L. (1927a). A law of comparative judgment. *Psychological Review*, **34**, 273-286.
- Thurstone, L. L. (1927b). Psychophysical analysis. *The American Journal of Psychology*, **38**, 368-389.
- Thurstone, L. L. (1928). The absolute zero in intelligence measurement. *Psychological Review*, **35**, 175-197.
- Thurstone, L. L., & Ackerson, L. (1929). The mental growth curve for the Binet tests. *Journal of Educational Psychology*, **20**, 569-583.
- Thurstone, L. L. (1931). The measurement of social attitudes. *Journal of Abnormal and Social Psychology*, **26**, 249-269.
- Torgerson, W. (1952). Multidimensional scaling : I. Theory and method. *Psychometrika*, **17**, 401-419.
- Torgerson, W. S. (1958). *Theory and Methods of Scaling*. New York: Wiley.
- Tucker, L. R. and Messick, S. (1963). An individual differences model for multidimensional scaling. *Psychometrika*, **28**, 333-367.
- 柳井晴夫 (1994). 多変量データ解析法 朝倉書店
- Young, F. W. (1975). An asymmetric Euclidean model for multi-process asymmetric data. *Paper presented at U.S.-Japan Seminar on MDS*, San Diego, U.S.A.
- Young, F. W., & Lewycky, R. (1979). *ALSCAL · 4 User's Guide*. Carrboro: Data Analysis and Theory Associates.
- Young, F. W., & Torgerson, W. S. (1967). TORSCA, a Fortran IV program for Shepard-Kruskal multidimensional scaling analysis. *Behavioral Science*, **12**, 498.
- Young, G., & Householder, A. S. (1938). Discussion of a set of points in terms of their mutual distances. *Psychometrika*, **3**, 19-22.
- Wilkinson, J. H. (1965). *The Algebraic Eigenvalue Problem*. Oxford: Oxford University Press.
- Zinnes, J. L., & Griggs, R. A. (1974). Probabilistic multidimensional unfolding analysis. *Psychometrika*, **39**, 327-350.

## 付録 A 演習問題の証明

ここでは、「多次元尺度構成法」における演習問題の証明を行なう。

演習 1.1 (証明略)

演習 1.2 (証明略)

演習 1.3 (定理 1 (Young-Householder) の証明)

点間距離がユークリッド距離であるための必要条件

(点間距離がユークリッド距離  $\Rightarrow B$  が p.s.d.) (証明)  $d_{ij}$  が (原点を対象  $n$  とする) ユークリッド距離ならば、( $\mathbf{a}_i = (a_{i1}, \dots, a_{ir})$  なる行ベクトルを考えると) 内積行列

$$B = \{b_{ij} = \phi(\mathbf{a}_i, \mathbf{a}_j) = \mathbf{a}_i \mathbf{a}_j^t\}, \quad (\text{A.1})$$

について

$$B = \left\{ \frac{1}{2}(d_{in}^2 + d_{jn}^2 - d_{ij}^2) \right\}, \quad (\text{A.2})$$

が成り立ち、行列  $B$  は  $n-1$  次対称行列で、かつ  $B = AA^t$  が成り立つ。この時、 $B$  は p.s.d. (正の半定符号) である。

(註 1) Y-H では、 $B$  の p.s.d. についての証明なし。(証明 1) 内積  $b_{ij} = \mathbf{a}_i \mathbf{a}_j^t$  から誘導される二次形式  $b_{ii} = \mathbf{a}_i \mathbf{a}_i^t$  によるノルムにより、ユークリッド距離が定義される。つまり、 $b_{ii} \geq 0$  より、 $B$  は p.s.d. (証明 2)  $B = AA^t$  の右辺の  $A$  に  $A$  の特異値分解の結果を代入。

点間距離がユークリッド距離であるための十分条件

( $B$  が p.s.d.  $\Rightarrow$  点間距離はユークリッド距離) (証明)  $B$  が p.s.d. なので、 $B = O\Delta O^t$  なる直交行列  $O$  が存在する (ここで、 $B$  は p.s.d. なので  $\Delta$  の対角要素はすべて正かゼロ)。つまり

$$\Delta = \text{diag}(\delta_1, \dots, \delta_r, 0, \dots, 0). \quad (\text{A.3})$$

そこで、 $D = \text{diag}(\sqrt{\delta_1}, \dots, \sqrt{\delta_r}, 0, \dots, 0)$  と置くと、

$$B = (OD)(OD)^t = AA^t. \quad (\text{A.4})$$

(註 2) ここでは、必要条件の時のような内積の性質は最初からは使えないので、 $B$  が p.s.d. との仮定から、行列の変換の知見を用いて、(A.2) 式の行列  $B$  (の要素) が内積表現できることを示すことによって、点間距離がユークリッド距離に相当することを証明している点が、重要。

演習 2.1 (定理 2 (Chino-Shiraiwa) の証明)

点間距離がヒルベルト距離であるための必要条件

(点間距離がヒルベルト距離  $\Rightarrow H$  が p.s.d.)

(証明)  $d_{jk} = \|\mathbf{v}_j - \mathbf{v}_k\|$ 、 $\bar{d}_{jk} = \|\mathbf{v}_j - i\mathbf{v}_k\|$  がヒルベルト距離ならば、 $(\mathbf{v}_j = (v_{j1}, \dots, v_{jr}))$  なる行ベクトルを考えると) エルミート内積行列

$$\mathbf{H} = \{h_{jk} = \phi(\mathbf{v}_j, \mathbf{v}_k) = \mathbf{v}_j \mathbf{\Lambda} \mathbf{v}_k^*\}, \quad (\text{A.5})$$

について、

$$\mathbf{H} = \left\{ \frac{1}{2}(d_{j0}^2 + d_{k0}^2 - d_{jk}^2) + \frac{1}{2}i(d_{j0}^2 + d_{k0}^2 - \bar{d}_{jk}^2) \right\}, \quad (\text{A.6})$$

が成り立ち、 $\mathbf{H}$  は  $n$  次のエルミート行列で、かつ  $\mathbf{H} = \mathbf{U}_1 \mathbf{\Lambda} \mathbf{U}_1^*$  が成り立つ。この時、 $\mathbf{H}$  は p.s.d.。

(註3) C-S でも証明はなし。(証明1) エルミート内積  $h_{jk} = \mathbf{v}_j \mathbf{\Lambda} \mathbf{v}_k^*$  から誘導される(エルミート形式  $h_{jj} = \mathbf{v}_j \mathbf{\Lambda} \mathbf{v}_j^*$  による) ノルムによる距離が定義される、つまり  $h_{jj} \geq 0$  より、 $\mathbf{H}$  は p.s.d.。(証明2)  $\mathbf{H} = \mathbf{U}_1 \mathbf{\Lambda} \mathbf{U}_1^* = \mathbf{P}_1 \mathbf{P}_1^*$  の右辺の  $\mathbf{P}_1$  に  $\mathbf{P}_1$  の特異値分解の結果を代入。

点間距離がヒルベルト距離であるための十分条件

( $\mathbf{H}$  が p.s.d.  $\Rightarrow$  点間距離はヒルベルト距離) (証明)  $\mathbf{H}$  が p.s.d. なので、 $\mathbf{H} = \mathbf{U} \mathbf{\Lambda}_0 \mathbf{U}^*$  なるユニタリ行列  $\mathbf{U}$  が存在する(ここで、 $\mathbf{H}$  が p.s.d. より、 $\mathbf{\Lambda}_0$  の対角要素は、すべて正またはゼロ) つまり、

$$\mathbf{\Lambda}_0 = \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0). \quad (\text{A.7})$$

そこで、 $\mathbf{F} = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}, 0, \dots, 0)$  と置くと、

$$\mathbf{H} = (\mathbf{U}\mathbf{F})(\mathbf{U}\mathbf{F})^* = \mathbf{P}\mathbf{P}^* \quad (\text{A.8})$$

(註4) ここでは、必要条件のようなエルミート内積の性質は最初からは使えないので、 $\mathbf{H}$  の p.s.d. の仮定から、行列の変換の知見のみを用いて、(A.6) 式の行列  $\mathbf{H}$  (の要素) が、エルミート内積表現できることを示すことにより、点間距離がヒルベルト距離に相当することを証明できる。実際には、別証明をしている。