The Society for Economic Studies
The University of Kitakyushu
Working Paper Series No.2017-4
(accepted in March 30, 2018)

コレスポンデンス分析の幾何学的意味と導出

齋藤朗宏
(北九州市立大学経済学部)

目的

分割表からカテゴリ間の関係を可視化する手法であるコレスポンデンス分析の応用範囲は広い．近年では，テキストマイニングの有力ソフトウェアである KH Coder(樋口，2004) に搭載されていることもあり，テキストデータの分析において単語間の関係を見るのに頻繁に用いられている．実際の応用場面としては，マーケティングにおける商品のポジショニングなどで使用されるケースは多い．

この手法を解説した書籍等も少なくはなく，近年ではたとえば Clausen(2015) などがある．ただ，その幾何学的な意味，明快な導出まで踏み込んだ解説というのは中々見られないようだ．そこで本稿では，コレスポンデンス分析の幾何学的な意味を確認した上でなるべく直感的にわかりやすい導出を試みる．

まず，行列計算の基礎的な部分について，その幾何学的な意味を確認し，その上でコレスポンデンス分析の意味について検討する．さらに，最もオーソドックスな手法と思われる共分散行列の固有値分解で，なぜプロットが可能なのかについて確認する．

行列計算の基礎

行列とベクトルの計算の意味

ベクトルに行列を作用させるということの意味を考える．たとえば，以下のような行列とベクトルを考える．

$$\boldsymbol{x} = \begin{bmatrix} 1 & 1 \end{bmatrix}', \quad \boldsymbol{A} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} 2 & 0.5 \\ 1 & 1.5 \end{bmatrix} \tag{1}$$

ここで，ベクトル $\boldsymbol{x}$ に対角行列 $\boldsymbol{A}$ を作用させると以下の通りである．

$$\boldsymbol{Ax} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \times 1 + 0 \times 1 \\ 0 \times 1 + 3 \times 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \tag{2}$$

これは，座標平面で考えれば，座標 $(1, 1)$ を座標 $(2, 3)$ に変換していることに相当する．仮に各要素の値が 2 であるベクトルに $\boldsymbol{A}$ を作用させるのであれば，その結果は $(4, 6)$ になる．つまり対角

行列 $\boldsymbol{A}$ を作用させるということは，座標平面上で，x 軸の値を 2 倍，y 軸の値を 3 倍することに相当し，ベクトルを，対角要素の倍率で引き延ばす．言い方を変えると，その座標を平面そのものの縮尺を，それぞれの軸の方向に変更すると考えることができる．

これを確認するために，座標平面上にプロットする．見やすさのために，行列 $\boldsymbol{B}$ を用いる．

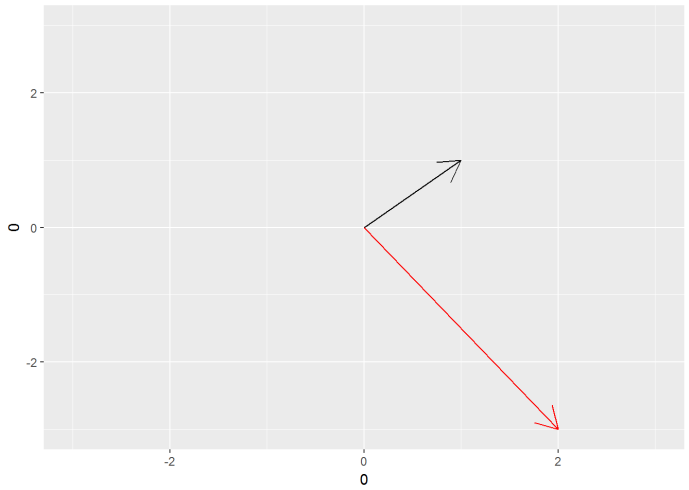

図 1：ベクトルの対角行列による変換

ここで，行列を作用させた結果は赤矢印である．つまり，黒矢印に相当する座標空間が赤矢印に相当する座標空間に置き換えられたと言える．5 × 5 の 25 個の座標を考え，それらをすべて変換すると図 2 のようになる．

ここで重要なのは，変換前の x 軸の値が等しいもの同士は，変換後の x 軸の値もかわらず，プロットした結果がどちらも長方形になっているという点である．つまり，x 軸の値は y 軸の値と無関係に決定され，逆に y 軸の値は x 軸の値と無関係に決定されるのである．これは，別の言い方をすれば，x 軸と y 軸は変換後も独立であり，直交しているという関係が維持されるということになる．

では，対角行列ではない場合にはどうなるのか，次にそれを確認する．

$$\boldsymbol{Cx} = \begin{bmatrix} 2 & 1.5 \\ 1 & 1.5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \times 1 + 1.5 \times 1 \\ 1 \times 1 + 1.5 \times 1 \end{bmatrix} = \begin{bmatrix} 3.5 \\ 2.5 \end{bmatrix} \tag{3}$$

たとえばこのようになり，その変換結果は図 3 の通りである．

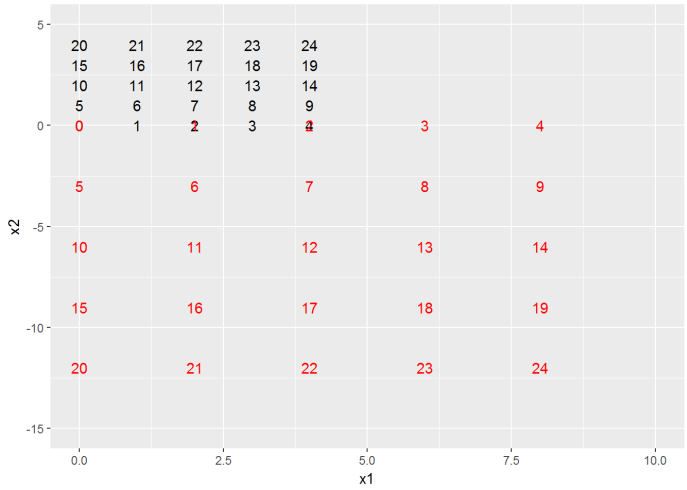

図 2：25 個の座標の変換 (1)

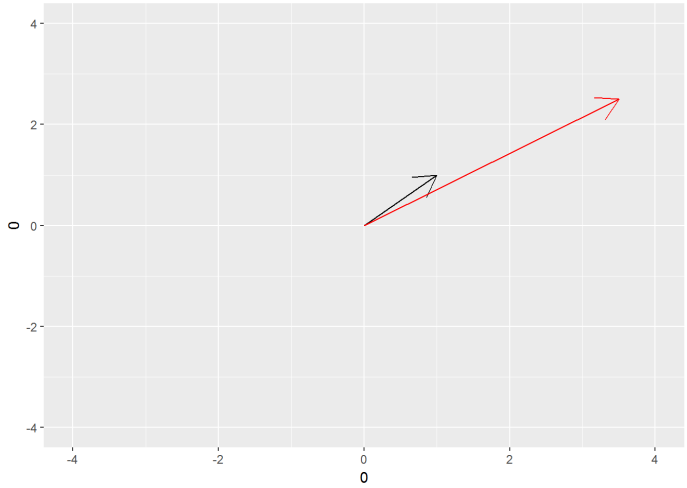

図 3：ベクトルの行列による変換

黒のベクトルが赤のベクトルに置き換えられたという意味では一見同じである．ただ，25 個置き換えると違いが確認できる．変換結果は図 4 の通りである．

0 から 4 までの並びに着目すると，黒，つまり変換前の場合には，y 軸の値は 0 で，x 軸の値が 0 から 4 まで変化している一方赤，つまり変換後の場合には，$(0,0)$ から $(8,4)$ と値が変化している．つまりこれは，x 軸が傾き $1/2$ の直線 (赤い点線) に置き換えられたことに相当する．同様に，y 軸は傾き 1 の直線に置き換えられている．このように，非対角行列の場合には，x 軸の値が y 軸に，y 軸の値が x 軸に影響を与えるため，単なる軸の引き延ばしや縮小だけではなく，軸同士の関係も置き換わることになる．これがベクトルに行列を作用させることの大まかな意味である．

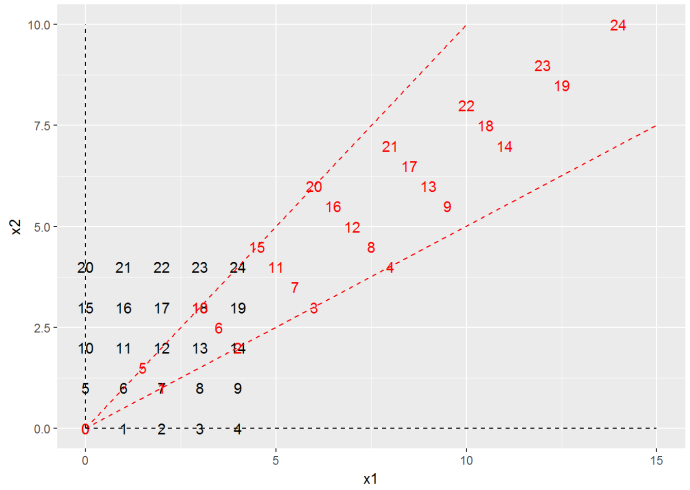

図 4：25 個の座標の変換 (2)

直交回転

さて，ベクトルに行列を作用させると以上のように軸の交わる角が変化したり，それぞれの軸の引き延ばし，縮小が行われた．ただ，今回説明するコレスポンデンス分析の場合，原則として軸同士の交わる角が変化するのは望ましくない (勿論，因子分析における斜交回転のように，軸同士の交わる角を変化させる回転は珍しくない)．尚，軸の引き延ばし，縮小については，変換後の値を定数倍することと大きな違いはなく，また，むしろ布置を見やすくするために必要なことも多いので，元々の値にこだわる必要はなく，使い方等から柔軟に考えればよい．

こういった場合には，軸同士の関係を変化させない変換のための行列が必要である．こういった行列のことを一般的に直交行列 (正確に言うと，行列から行，列を取り出したとき，行同士，列同士が直交し，ベクトルの長さが 1 であるというのが定義になる．) と呼び，直交行列を用いて座標を回転させることを直交回転と呼ぶ．2 次元の直交行列の一例は以下の通りである．

$$\boldsymbol{T} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{4}$$

これは，座標を角 θ だけ回転させる行列である．実際，点 (x_1, y_1) はどう変化するだろうか．これは，ベクトル $[x_1, y_1]$ を角 θ だけ回転させると考えた方がわかりやすい．たとえば以下のような移動である．ここでは，簡単のため長さを 1 としている．

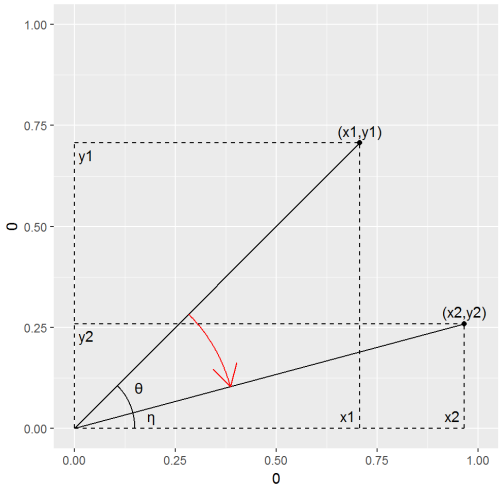

図 5：直交回転の例

ベクトル $[x_1, y_1]$ の長さは 1 なので，以下が容易に導かれる．

$$x_1 = \cos(\theta + \eta) = \cos\theta\cos\eta - \sin\theta\sin\eta \tag{5}$$
$$y_1 = \sin(\theta + \eta) = \sin\theta\cos\eta + \cos\theta\sin\eta \tag{6}$$

回転後のベクトル $[x_2, y_2]$ の長さも当然 1 なので，以下となる．

$$x_2 = \cos\eta \tag{7}$$
$$y_2 = \sin\eta \tag{8}$$

2 次正方行列 $\boldsymbol{T}$ とのかけ算でベクトルを回転させると言うことは，

$$\boldsymbol{T}\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} t_{11}x_1 + t_{12}y_1 \\ t_{21}x_1 + t_{22}y_1 \end{bmatrix} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \tag{9}$$

が成立するということで，つまり

$$x_2 = t_{11}x_1 + t_{12}y_1 \tag{10}$$

である．ここから，

$$x_2 = t_{11}(\cos\theta\cos\eta - \sin\theta\sin\eta) + t_{12}(\sin\theta\cos\eta + \cos\theta\sin\eta) \tag{11}$$

となることがわかる．$\cos\eta$ は括弧の中のそれぞれ第 1 項に入っていて，それぞれ $\sin\theta, \cos\theta$ がかかっている．そこで，$(\sin^2\theta + \cos^2\theta)\cos\eta$ の形になるように調整する．その際，括弧の中の第 2 項が消えるように適切な係数を探すと，$t_{11} = \cos\theta$，$t_{12} = \sin\theta$ が適切である．それぞれ代入すると，

$$x_2 = \cos\theta(\cos\theta\cos\eta - \sin\theta\sin\eta) + \sin\theta(\sin\theta\cos\eta + \cos\theta\sin\eta) \tag{12}$$
$$= \cos^2\theta\cos\eta - \cos\theta\sin\theta\sin\eta + \sin^2\theta\cos\eta + \cos\theta\sin\theta\sin\eta \tag{13}$$
$$= \cos^2\theta\cos\eta + \sin^2\theta\cos\eta = (\sin^2\theta + \cos^2\theta)\cos\eta = \cos\eta \tag{14}$$

となり，$x_2 = \cos\eta$ が成立する変換となっていることが確認できる．

$$y_2 = t_{21}x_1 + t_{22}y_1 \tag{15}$$

も同様に考えるといい．

$$y_2 = t_{21}(\cos\theta\cos\eta - \sin\theta\sin\eta) + t_{22}(\sin\theta\cos\eta + \cos\theta\sin\eta) \tag{16}$$

であり，ここで $t_{21} = -\sin\theta$，$t_{22} = \cos\theta$ とすると，

$$\begin{aligned} y_2 &= -\sin\theta(\cos\theta\cos\eta - \sin\theta\sin\eta) + \cos\theta(\sin\theta\cos\eta + \cos\theta\sin\eta) && (17)\\ &= -\sin\theta\cos\theta\cos\eta + \sin^2\theta\sin\eta + \sin\theta\cos\theta\cos\eta + \cos^2\theta\sin\eta && (18)\\ &= \sin^2\theta\sin\eta + \cos^2\theta\sin\eta = (\sin^2\theta + \cos^2\theta)\sin\eta = \sin\eta && (19) \end{aligned}$$

となり，$y_2 = \sin\eta$ が成立する変換となっていることが確認できる．以上から $\boldsymbol{T}$ は 2 次元平面において座標を角 θ だけ回転させる意味を持つ行列になっている．

尚，2 つ以上のベクトルに行列を作用させる計算が行列と行列のかけ算と考えることができ，たとば行列 $\boldsymbol{W}$ に 2 つのベクトル $\boldsymbol{xy}$ を作用させる計算は

$$\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} = \begin{bmatrix} w_{11}x_1 + w_{12}x_2 & w_{11}y_1 + w_{12}y_2 \\ w_{21}x_1 + w_{22}x_2 & w_{21}y_1 + w_{22}y_2 \end{bmatrix} \tag{20}$$

となり，1 列目がベクトル $\boldsymbol{x}$，2 列目がベクトル $\boldsymbol{y}$ の変換結果である．また，計算上，行列とベクトルの順番が入れ替わることもあり得る．この場合，たとえば以下のように考える．

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix}' \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}' = \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} \begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \end{bmatrix} = \begin{bmatrix} w_{11}x_1 + w_{12}x_2 & w_{21}x_1 + w_{22}x_2 \\ w_{11}y_1 + w_{12}y_2 & w_{21}y_1 + w_{22}y_2 \end{bmatrix} \tag{21}$$

これは，先ほどの変換結果の転置であるので，以下の関係が成立していることがわかる．

$$\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} = \left[\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix}' \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}' \right]', \quad \boldsymbol{WX} = (\boldsymbol{X}'\boldsymbol{W}')' \tag{22}$$

$$\left[\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} \right]' = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix}' \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}', \quad (\boldsymbol{WX})' = \boldsymbol{X}'\boldsymbol{W}' \tag{23}$$

一般的に，データ行列は行にオブザベーション，列に変数を配置することが多い．そのため，それぞれの変数を利用してデータを変換，新しい変数を作成するような計算を行う場合には，$\boldsymbol{x}$，$\boldsymbol{y}$ がオブザベーションとなり，それに係数をかけることで新しい変数の値がそれぞれのオブザベーションに対して求まるため，かけ算の左側にデータを，右側に係数を置く形は都合がよい．

コレスポンデンス分析の幾何学的意味

コレスポンデンス分析の例

表 1 は，行にテレビの好きなジャンル，列にテレビを購入する際に重視する機能を配置した分割表である．全体的にバラエティを好きな人数が多いため，たとえば録画機能を重視する人の中で，一番多いのはバラエティを好む人だからといって，バラエティ好きは録画機能を好むとは言えない．こういった場合に，カテゴリ間の関係を見やすく表現するのがコレスポンデンス分析である．このデータからは，たとえば図 6 のような結果が得られる．

表 1：好きなジャンルと重視する機能

	価格	画面サイズ	録画機能
アニメ	16	20	22
バラエティ	37	40	37
映画	9	13	10

図 6 からは，バラエティを好む人はむしろ価格を重視し，録画機能への関心が低いことが見て取れる．

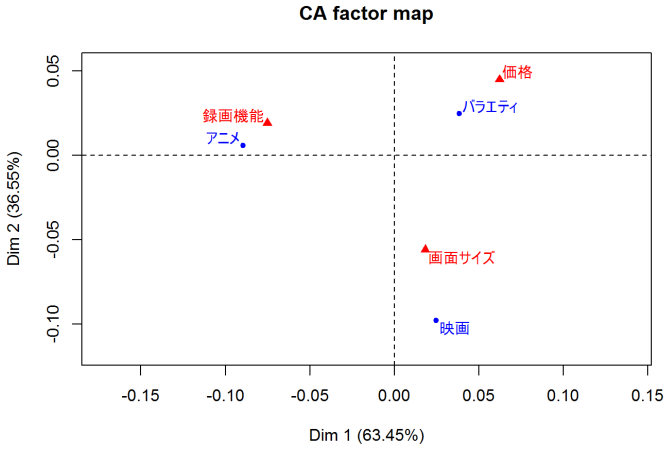

図 6：コレスポンデンスの例

さて，このような関係はどのようにして得られるだろうか．それを検討するにあたり，まず統計学における距離について確認する．

様々な距離の定義

座標空間上の距離を考える．たとえば $x, y2$ 次元の平面において，2 点 $a = (1, 2)$ と $b = (3, 5)$ 間の距離を計算する．

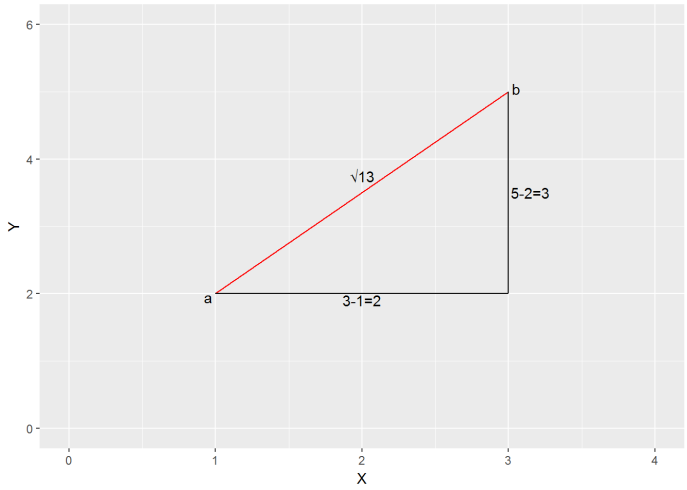

図 7：2 点間の距離

これは，上図の赤線の長さを求めることになり，直角三角形の斜辺の長さを求める問題であるため，三平方の定理で容易に求められる．

$$d_{ab} = \sqrt{(b_x - a_x)^2 + (b_y - a_y)^2} + \sqrt{(3-1)^2 + (5-2)^2} = \sqrt{13} \tag{24}$$

より一般的に，n 次元の座標空間における距離は，以下のように定義される．

$$d_{ab} = \sqrt{\sum_{i=1}^{n} (b_i - a_i)^2} \tag{25}$$

これは，ユークリッド距離と呼ばれるものである．ただ，このユークリッド距離は，分割表に用いるのは現実的ではない．そのことを確認するために，表 2 のような極端な数値例を考える．

表 2：数値例

	A	B	C
X	1	9	80
Y	2	10	88
Z	9	2	89

ここで，A と C は小さな順に XYZ となっている一方，B は ZYX である．A と B でが大小関係がほぼ逆転している一方，A と C とでは順が一致している点など，直感的には，A と B よりは A と C の方が近い関係に見える．しかしここで，A と B，A と C のユークリッド距離はそれぞれ以下のように求められる．

$$d_{AB} = \sqrt{(19-1)^2 + (10-2)^2 + (2-9)^2} = \sqrt{324 + 64 + 49} = 20.9 \tag{26}$$

$$d_{AC} = \sqrt{(80-1)^2 + (88-2)^2 + (89-9)^2} = \sqrt{6241 + 7396 + 6400} = 141.55 \tag{27}$$

ここから，A と B の方がユークリッド距離では近くなっていることが確認できる．これは，B に該当するのが A の約 2.6 倍である一方，C に該当する人数は 21.4 倍であり，単純に該当する人数の

差を用いたのではそもそもの人数差の影響を大きく受けてしまうということである．そこで，A，B，C を選択した人数で割った割合で考えると以下のようになる．

$$d_{AB} = \sqrt{(0.61 - 0.08)^2 + (0.32 - 0.17)^2 + (0.06 - 0.75)^2} = \sqrt{0.2809 + 0.0225 + 0.4761} = 0.88 \tag{28}$$

$$d_{AC} = \sqrt{(0.31 - 0.08)^2 + (0.34 - 0.17)^2 + (0.35 - 0.75)^2} = \sqrt{0.0529 + 0.0289 + 0.16} = 0.49 \tag{29}$$

こうすると，直感的には問題ないように見える．さて，ここで，X と Y，Y と Z 間の距離を考える．つまり，以下である．

$$d_{XY} = \sqrt{(0.02 - 0.01)^2 + (0.1 - 0.19)^2 + (0.88 - 0.8)^2} = \sqrt{009 + 0.0081 + 0.0064} = 0.12 \tag{30}$$

$$d_{YZ} = \sqrt{(0.09 - 0.02)^2 + (0.02 - 0.1)^2 + (0.89 - 0.88)^2} = \sqrt{0.0049 + 0.0064 + 0} = 0.11 \tag{31}$$

以上のように，YZ 間の距離よりも XY 間の距離の方が大きいということができる．この結果は適切だろうか．X と Y を比較すると，A と C は Y の方がそれぞれ 2 倍，1.1 倍大きく，B は X の方が 1.9 倍大きい．一方 Y と Z では，A は Z が 4.5 倍大きく、B では Y が 5 倍，C は Z が 1.01 倍大きい．また順序関係で見ても，X と Y は小さい順に ABC だが，Z は BAC である．

これらを見比べると，A の 4.5 倍差や B の 5 倍差，また順序関係の違いっも大きく YZ の方が違いは大きくも見える．しかしそうはなっていない．計算を見ると，C における X と Y は 1.1 倍 Y の方が大きいが，全体に占める割合が大きいため，Y と X の差を取ると 0.8 という大きな値になっており，これは，A における Y と Z を考えると，Z が 4.5 倍大きいにもかかわらず差は 0.7 となっている点よりも，全体に与える影響が大きいことがわかる．

このように，そもそもの人数が他と比べて飛び抜けて多い C における差が，XY の方が大きく，結果にはその違いが大きく反映されるため，そもそもの人数の少ない A や B での差を覆い隠してしまっていると考えることができる．

ここからも，合計値で割る方法では，今度は XYZ それぞれの中での値の違いの影響が出てしまうということがわかる．そこで，その影響を補正するために，さらに ABC それぞれの比率で補正した距離は χ^2 距離と呼ばれ，以下のように定義される．ここで，r_i は，i 番目の行が全体に占める割合である．

$$c_{ab} = \sqrt{\sum_{i=1}^{n} \frac{(b_i - a_i)^2}{r_i}} \tag{32}$$

たとえば以下の通りである．

$$c_{XY} = \sqrt{\frac{(0.02-0.01)^2}{0.04} + \frac{(0.1-0.19)^2}{0.1} + \frac{(0.88-0.8)^2}{0.86}} = \sqrt{0.0025+0.081+0.0074} = 0.3 \tag{33}$$

$$c_{YZ} = \sqrt{\frac{(0.09-0.02)^2}{0.04} + \frac{(0.02-0.1)^2}{0.1} + \frac{(0.89-0.88)^2}{0.86}} = \sqrt{0.1225+0.064+0} = 0.43 \tag{34}$$

$$\tag{35}$$

ここから，χ^2 距離では YZ の方が若干離れている．これは，直感とも合致する．

最後に，式から χ^2 距離の意味を確認する．仮に，A の合計人数を a，B の合計人数を b，また，X の合計人数を x のように置き，X かつ A の人数を (xa)，全体の合計人数を t と置くと，表 3 のようになる．

表 3：数値例

	A	B	C	合計
X	(xa)	(xb)	(xc)	x
Y	(ya)	(yb)	(yc)	y
Z	(za)	(zb)	(zc)	z
合計	a	b	c	t

ここで，単純に人数を使った場合の A と B のユークリッド距離は以下の通りである．

$$d_{AB} = \sqrt{(xb-xa)^2 + (yb-ya)^2 + (zb-za)^2} \tag{36}$$

それに対して，ABC それぞれの合計数で割った場合のユークリッド距離は以下のようになる．

$$d_{AB} = \sqrt{\left(\frac{xb}{b} - \frac{xa}{a}\right)^2 + \left(\frac{yb}{b} - \frac{ya}{a}\right)^2 + \left(\frac{zb}{b} - \frac{za}{a}\right)^2} \tag{37}$$

最後に，χ^2 距離は以下のようになる．

$$c_{AB} = \sqrt{\frac{\left(\frac{xb}{b} - \frac{xa}{a}\right)^2}{\frac{x}{t}} + \frac{\left(\frac{yb}{b} - \frac{ya}{a}\right)^2}{\frac{y}{t}} + \frac{\left(\frac{zb}{b} - \frac{za}{a}\right)^2}{\frac{z}{t}}} \tag{38}$$

2 乗の部分があるため完全に一致はしないが，χ^2 距離の場合には xa に対してそれぞれの行，列の合計数である x と a で割る形になっており，これにより，xa から，それぞれの合計数の効果が除去されていることが確認できる．

位置関係の図示

χ^2 距離を用いることで，分割表からカテゴリ間の距離を求めることができた．カテゴリ間の距離がわかれば，その距離を利用して，各カテゴリの位置関係を図示することができる．テレビの機能におけるカテゴリ間の χ^2 距離は表 4 の通りとなった．

表 4 : 4 カテゴリ間の距離

	ブランド	価格	画面サイズ	録画機能
ブランド	0.00	0.23	0.28	0.32
価格	0.23	0.00	0.18	0.23
画面サイズ	0.28	0.18	0.00	0.17
録画機能	0.32	0.23	0.17	0.00

先ほどのデータ例に，カテゴリ「ブランド」が加わっている．

当然のことながら，たとえばブランドとブランドとの差は 0 だし，ブランドと価格の距離と，価格とブランドの距離は定義上等しい．ただし，データの性質によっては (たとえば，A さんから見た B さんの親しみやすさと，B さんから見た A さんの親しみやすさなど) この性質は成立しないので，用いるデータに，この対称性があるのか確認するのは重要である．

さて，ここで，ブランドと録画機能との距離に注目する．この距離は.32 である．この距離を図示すると，たとえば図 8 のようになる．

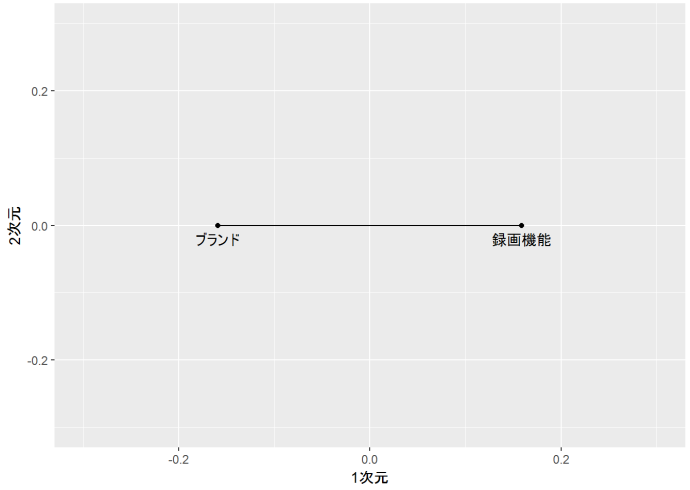

図 8 : 2 点間の位置関係

このように，2 点間の距離は x 軸上で表現できる．別の言い方をすれば，1 次元上で表現可能と言える．ではここに，画面サイズを付け加え，3 点間の距離を考えたらどうなるだろうか．仮に，ブランドと画面サイズ，録画機能と画面サイズとの間の距離が等しく 0.16 であるならば簡単で，以下のような一直線上に来る．

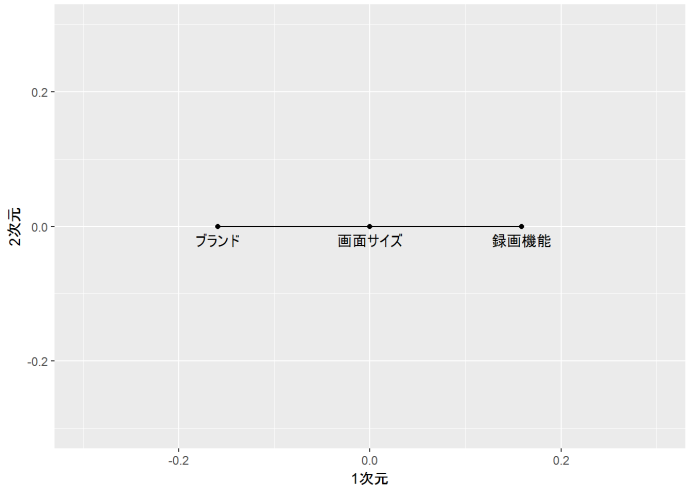

図 9：1 次元で収まる場合の 3 点間の位置関係

この図は，ブランドからも録画機能からも画面サイズは等しく 0.16 だけ離れている点を確認せよ．

さて，実際には，ブランドと画面サイズは 0.28，録画機能と画面サイズとは 0.17 だけ離れている．この場合，この 3 点間の距離は以下のように求めることができる．

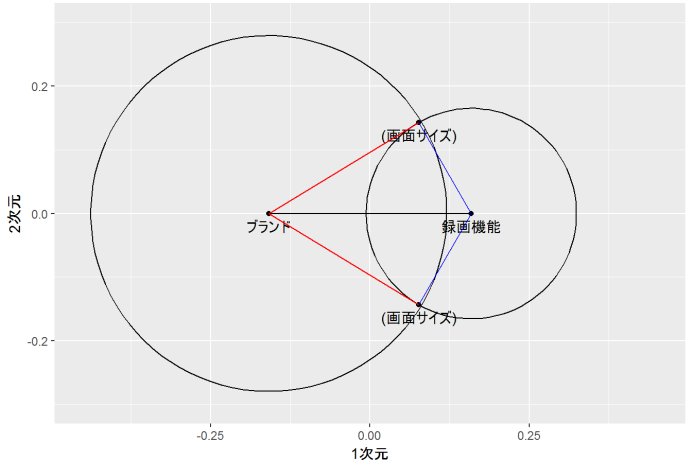

図 10：3 点間の位置関係

ブランドと録画機能との間の距離は 0.28 なので，ブランドの座標を中心に半径 0.28 の円を描いたときに，その円周上のいずれかに画面サイズの座標は定まることになる．同様に，録画機能と画面サイズとの間の距離が 0.17 であることから，録画機能の座標を中心とした半径 0.28 の円周上に画面サイズの座標が存在することになり，その両方を満たす円の交点 2 カ所が画面サイズの座標になり得る場所と言える．

ここから 2 つのことがわかる．1 つめは，対象間の距離を正しく表現するために必要な次元数である．2 つの対象の間の距離を正しく表現するためには 1 次元の直線があればいいが，3 つの対象の間の距離を表現するためには 2 次元の平面が必要である．今回の図に価格を加え 4 つの対象を

考えると，たとえばブランドの座標を中心として，半径がブランドと価格の距離 0.23 である球を描き，同じように録画機能，画面サイズからも球を描き，3 つの球がすべて交わる点である必要があり，つまり 3 次元空間が必要になる．以下同様に，n 個の対象の距離関係を正しく図示するためには，$n-1$ 次元空間が必要になる．

2 つ目は，座標を配置する位置である．ブランドと録画機能の座標が定まった状態でも価格の座標は 1 点に定まらなかった．実際には，ブランドや録画機能の座標もこの 2 点に必ずしも定まるものではないため，取りうる座標の組み合わせは無数にある．一般的には，重心の座標が原点になるように定めることも多いが，それにしても，原点をを中心に座標を回転させれば，取りうる値は無数に存在する．

次元縮約

次に，先程述べた 2 つの点が持つ意味について検討する．距離がわかっている状態で対象を空間上に配置する場合，n 個の対象の距離関係を正しく図示するためには，$n-1$ 次元空間が必要になる．配置する対象が 3 つ以下であれば，その関係を図示するのは，前節で確認したとおり難しくない．対象が 4 つある場合，3 次元の空間に配置することになり，それは不可能ではないが，第一に直感的にわかりづらい．第二にコンピュータ上に表示する分には見やすくすることもできるが，印刷する場合見づらくなる．対象が 5 つを超えると，4 次元以上の空間が必要になり，人間には不可能である．

そこで，3 次元，あるいはそれ以上の空間に配置された対象の関係を，可能な限り維持したまま寄り少ない次元に落とす方法を考える．これは，次元縮約と呼ばれる．コレスポンデンス分析も含め，基本的には多次元の情報を 2 次元に落とすために使われるが，わかりやすさのために 2 次元の情報を 1 次元に縮約する方法を基本的に説明する．多次元の情報の縮約は，この方法を拡張することで自然に導かれる．

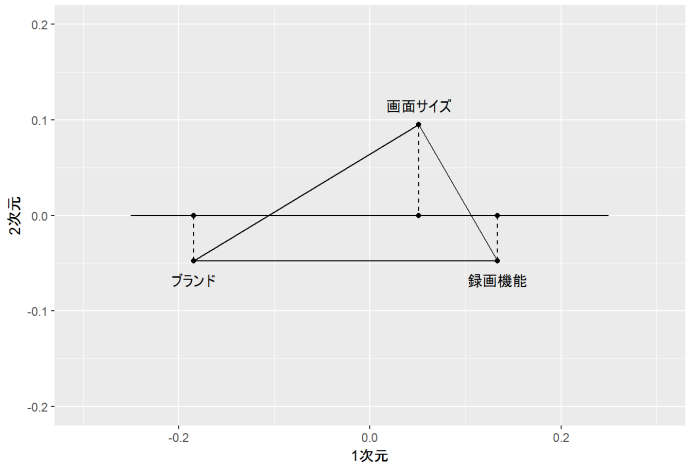

図 11：次元縮約の例 (1)

最も簡単に，2 次元平面の情報を 1 次元に落とし込む方法は，座標の一方の軸の値のみを使う方

法である．たとえばX軸の値のみを使うと，2次元の情報は図11のような形で1次元に落とされる．この図を見ると，確かにブランドと録画機能が大きく離れ，画面サイズはその間やや録画機能寄りという情報は正しく保存され，問題はないように見える．

では，図12の場合だとどうだろうか．

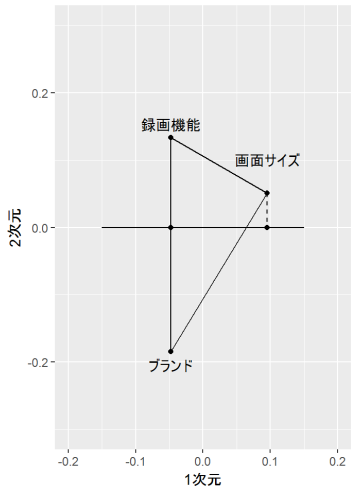

図12：次元縮約の例 (2)

これは，先ほどの図のX軸とY軸を入れ替えたものである．つまり，2次元平面における3点の関係は正しく維持されている．しかし，これを1次元に縮約すると，録画機能とブランドが同じ点になり，画面サイズだけが離れる結果となった．これは，2次元平面の情報が全く反映されていない結果である．

このように，次元を減らすという行為から得られる結果は，どのように回転させて見るかによって異なり，回転のパターンは無数にある．

3次元へのプロットと次元縮約

3次元を図示するのは難しいので，基本的な範囲にとどめるが，これまで述べてきた3点に，新たな点「価格」を追加することを考える．

表4から，点「ブランド」から距離0.23であり，かつ「録画機能」からも0.23,「画面サイズ」から0.18の点ということになる．これも同じように3つの点からそれぞれの距離の円を描くと図13の通りとなる．

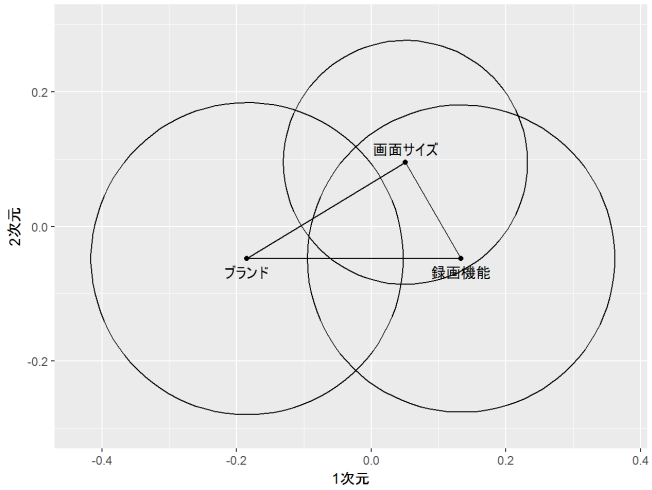

図 13：2 次元空間上での 4 点目の探索

距離関係が維持されるためには，3 つの円すべてが交わる点である必要があるが，2 次元平面上にはそのような点は存在しない．こういった場合には，たとえば点「ブランド」から半径 0.23 の球，を描き，他の 2 点からも同じように球を描き，3 つの球が交わる点を探すというように，4 つの距離関係を正しく表現するためには通常 3 次元の空間を用いる必要がある．実際に，3 次元空間上に 4 つ目の点を追加すると図 14 のようになる．

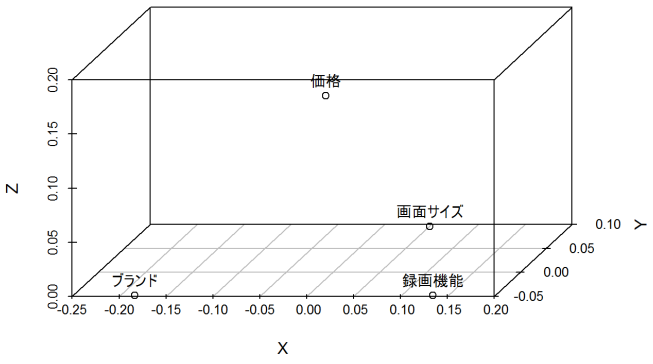

図 14：3 次元空間上への 4 点目の布置

紙という 2 次元平面上で 3 次元の位置関係を理解するのはやはり難しく，2 次元への次元縮約が求められる．5 つの点には 4 次元空間，6 つの点には 5 次元空間という用に，点の数 n に対して，正しく距離関係を維持するには通常 $n-1$ 次元が必要であり，そもそも 5 つ以上の点を正確に表示するのは不可能である．

2 次元→ 1 次元で行ったように，3 次元目，Z 軸を除いて X, Y2 つの軸に関わる座標のみに縮約する．これは，上の 3 次元散布図を真上からのぞき込んでいる状態に相当し，図 15 のようになる．

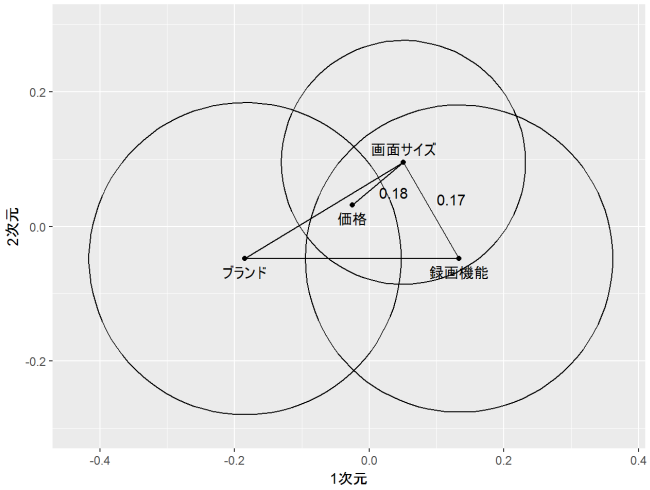

図 15：2 次元空間への次元縮約

3 つの円の重なり合う領域のほぼ中心に点「価格」は存在し，概ね位置関係の情報は正しく縮約されている．ただ，たとえば「画面サイズ」と「録画機能」の距離は 0.17 であり，「価格」と「画面サイズ」の距離は 0.18 である．しかし，この図では「価格」と「画面サイズ」の方が近く表示されている．このように，やはり単純に 1 次元分情報を削るという方法では，情報損失は避けられない．そのため，情報損失を最小限にとどめる回転が必要とされる．

回転の意味

最適なを考えるにあたり，統計学における分散と，ベクトルの長さの意味について考える．

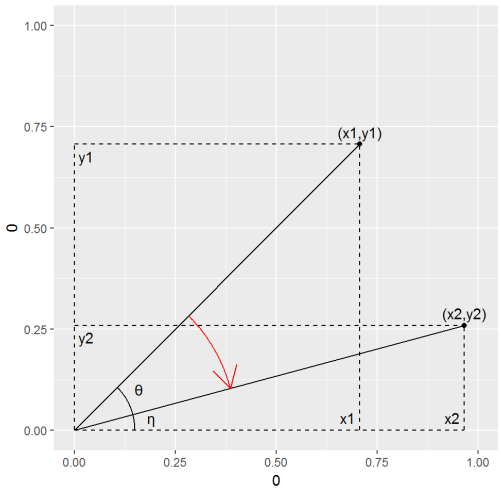

図 5(再掲)：直交回転の例

先ほども示した図 5 は，ベクトル $[x_1, y_1]$ を回転させて $[x_2, y_2]$ に変換したものである．ここで，先ほども示したとおり，それぞれの長さは 1 であり，

$$x_1 = \cos(\theta + \eta) \tag{39}$$
$$y_1 = \sin(\theta + \eta) \tag{40}$$
$$x_2 = \cos\eta \tag{41}$$
$$y_2 = \sin\eta \tag{42}$$

である．長さ 1 になるように設定したのだから当然であるが，

$$\sqrt{x_1^2 + y_1^2} = \sqrt{\cos^2(\theta + \eta) + \sin^2(\theta + \eta)} = 1 \tag{43}$$
$$\sqrt{x_2^2 + y_2^2} = \sqrt{\cos^2\eta + \sin^2\eta} = 1 \tag{44}$$

である．さて，ここで，この「2 乗の和」の意味について考える．この 2 乗した合計は何を意味しているだろうか．これは，幾何学的には原点からの距離ということになる．ここで，変数 x と y について，その平均値 (重心) を原点 0 と中心化すると，x_1^2，y_1^2 はそれぞれ平均からの差の 2 乗ということになり，複数の点が存在するとき，それぞれについて x_i^2 を求め，その平均を算出すると変数 x の分散が求められることになる．このように，分散とは，幾何学的には「重心から見た各軸における距離の平均」を意味していることがわかる．

また，たとえば $\theta + \eta = \pi/4$，$\eta = \pi/12$ とすると，以下のようになる．

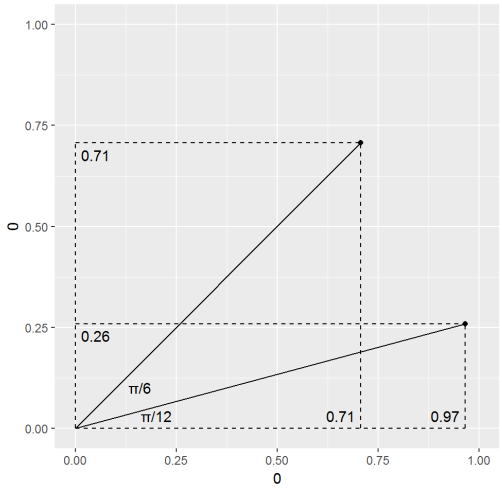

図 16：回転の数値例

回転前は $x_1^2 = 0.71^2 = 0.5$ である一方で，回転後は $x_2^2 = 0.97^2 = 0.93$ と，回転前後で変数 x における平均からの差の 2 乗の値は変化している．しかし，

$$x_1^2 + y_1^2 = 0.71^2 + 0.71^2 = 1 \tag{45}$$
$$x_2^2 + y_2^2 = 0.97^2 + 0.26^2 = 1 \tag{46}$$

のように，それぞれの変数の平均からの差の 2 乗を合計したものは変わらない．回転前は距離に占める割合が，x 軸と y 軸それぞれ半分ずつであったのが，回転後は x 軸が大半を占め，y 軸はわず

かになっている．つまり，回転というのは，平均からの偏差の 2 乗の合計を変えない範囲で，それぞれの変数が占める偏差の 2 乗の割合を調整する作業と見なすことができる．幾何学的に言えば原点からの距離を変えない範囲でそれぞれの軸が占める距離の割合を調整する作業となる．

さて，これを踏まえて，最適な回転とはどのようになされると考えればいいだろうか．それに対する一つの回答としては，表示される部分の分散を可能な限り大きくする方向を探すという方法が考えられる．

図 11 と 12 を思い出してみると，図 12 の 1 次元目は，図 11 の 1 次元目と比較して実態を反映していないと考えられる．実態が反映していないと感じる理由は，1 次元目で説明されない 2 次元目に，録画機能とブランドの間が大きく離れているという散らばり方の重要な情報が提示されているからである．このように，散らばりの大きい方向には重要な情報が含まれており，可能な限り多くの情報を含んだ次元縮約を行う際に，分散が大きくなる方向を探すというのは合理的である．そこで，分散が最大になる回転のしかたについて考える．

座標の再定義

これまで，まず χ^2 距離を算出し，そこから適当に何点か設定し，それを元に全体の座標を定める方法で説明してきたが，これは手間も多くあまり現実的な方法ではない．元々の分割表から直接的に座標を求めたい．

この問題は，つまり表 3 のような元の分割表を用いて AB 間の距離を考えるとユークリッド距離

$$d_{AB} = \sqrt{(xb - xa)^2 + (yb - ya)^2 + (zb - za)^2} \tag{47}$$

が出てきてしまうので,，同じ計算をしたときに χ^2 距離

$$c_{AB} = \sqrt{\frac{\left(\frac{xb}{b} - \frac{xa}{a}\right)^2}{\frac{x}{t}} + \frac{\left(\frac{yb}{b} - \frac{ya}{a}\right)^2}{\frac{y}{t}} + \frac{\left(\frac{zb}{b} - \frac{za}{a}\right)^2}{\frac{z}{t}}} \tag{48}$$

が出るように，行列の方を変換したいということである．統一性のために，距離は行間で考えるものとする場合，以下のようにすることでその変換はなされる．

$$\boldsymbol{A} = \begin{bmatrix} 1/(a/t) & 0 & 0 \\ 0 & 1/(b/t) & 0 \\ 0 & 0 & 1/(c/t) \end{bmatrix} \begin{bmatrix} (xa)/t & (xb)/t & (xc)/t \\ (ya)/t & (yb)/t & (yc)/t \\ (za)/t & (zb)/t & (zc)/t \end{bmatrix}' \begin{bmatrix} 1/\sqrt{x/t} & 0 & 0 \\ 0 & 1/\sqrt{y/t} & 0 \\ 0 & 0 & 1/\sqrt{z/t} \end{bmatrix} \tag{49}$$

$$= \begin{bmatrix} t/a & 0 & 0 \\ 0 & t/b & 0 \\ 0 & 0 & t/c \end{bmatrix} \begin{bmatrix} (xa)/t & (ya)/t & (za)/t \\ (xb)/t & (yb)/t & (zb)/t \\ (xc)/t & (yc)/t & (zc)/t \end{bmatrix} \begin{bmatrix} 1/\sqrt{x/t} & 0 & 0 \\ 0 & 1/\sqrt{y/t} & 0 \\ 0 & 0 & 1/\sqrt{z/t} \end{bmatrix} \tag{50}$$

つまり，分割表を度数ではなく割合に変換した上で，行間の距離の形になるように必要に応じて転置し，左から用いる行列の行周辺和の逆数を対角要素に持つ行列を，右側から用いる行列の列周辺和の平方根の逆数を対角要素に持つ行列を掛け合わせることでその変換はなされるということである．

この意味は，実際に続きを計算するとわかりやすい．

$$\boldsymbol{A} = \begin{bmatrix} (xa)t/ta & (ya)t/ta & (za)t/ta \\ (xb)t/tb & (yb)t/tb & (zb)t/tb \\ (xc)t/tc & (yc)t/tc & (zc)t/tc \end{bmatrix} \begin{bmatrix} 1/\sqrt{x/t} & 0 & 0 \\ 0 & 1/\sqrt{y/t} & 0 \\ 0 & 0 & 1/\sqrt{z/t} \end{bmatrix} \tag{51}$$

$$= \begin{bmatrix} (xa)t/ta\sqrt{x/t} & (ya)t/ta\sqrt{y/t} & (za)t/ta\sqrt{z/t} \\ (xb)t/tb\sqrt{x/t} & (yb)t/tb\sqrt{y/t} & (zb)t/tb\sqrt{z/t} \\ (xc)t/tc\sqrt{x/t} & (yc)t/tc\sqrt{y/t} & (zc)t/tc\sqrt{z/t} \end{bmatrix} = \begin{bmatrix} (xa)/a\sqrt{x/t} & (ya)/a\sqrt{y/t} & (za)/a\sqrt{z/t} \\ (xb)/b\sqrt{x/t} & (yb)/b\sqrt{y/t} & (zb)/b\sqrt{z/t} \\ (xc)/c\sqrt{x/t} & (yc)/c\sqrt{y/t} & (zc)/c\sqrt{z/t} \end{bmatrix} \tag{52}$$

ここで，AB 間の，つまり 1 行目と 2 行目の間のユークリッド距離を求めると

$$d_{AB} = \sqrt{\left(\frac{(xb)}{b\sqrt{x/t}} - \frac{(xa)}{a\sqrt{x/t}}\right)^2 + \left(\frac{(yb)}{b\sqrt{y/t}} - \frac{(ya)}{a\sqrt{y/t}}\right)^2 + \left(\frac{(zb)}{b\sqrt{z/t}} - \frac{(za)}{a\sqrt{z/t}}\right)^2} \tag{53}$$

$$= \sqrt{\left(\frac{1}{\sqrt{x/t}}\left(\frac{(xb)}{b} - \frac{(xa)}{a}\right)\right)^2 + \left(\frac{1}{\sqrt{y/t}}\left(\frac{(yb)}{b} - \frac{(ya)}{a}\right)\right)^2 + \left(\frac{1}{\sqrt{z/t}}\left(\frac{(zb)}{b} - \frac{(za)}{a}\right)\right)^2} \tag{54}$$

$$= \sqrt{\frac{\left(\frac{(xb)}{b} - \frac{(xa)}{a}\right)^2}{x/t} + \frac{\left(\frac{(yb)}{b} - \frac{(ya)}{a}\right)^2}{y/t} + \frac{\left(\frac{(zb)}{b} - \frac{(za)}{a}\right)^2}{z/t}} \tag{55}$$

となり，χ^2 距離と一致する．よって，この変換を用い，

$$A = \begin{bmatrix} (xa)/a\sqrt{x/t} & (ya)/a\sqrt{y/t} & (za)/a\sqrt{z/t} \end{bmatrix} \tag{56}$$

$$B = \begin{bmatrix} (xb)/b\sqrt{x/t} & (yb)/b\sqrt{y/t} & (zb)/b\sqrt{z/t} \end{bmatrix} \tag{57}$$

$$C = \begin{bmatrix} (xc)/c\sqrt{x/t} & (yc)/c\sqrt{y/t} & (zc)/c\sqrt{z/t} \end{bmatrix} \tag{58}$$

$$\tag{59}$$

と 3 点の座標を定めれば，3 点の間の距離は，χ^2 距離を正しく維持していることになり，布置として用いることができる．ただし，行列 $\boldsymbol{A}$ を転置しても，XYZ の座標として用いることはできないという点には注意が必要である．

YXZ の座標は以下の通りである．導出方法はほぼ同じなので省略する．

$$\boldsymbol{X} = \begin{bmatrix} t/x & 0 & 0 \\ 0 & t/y & 0 \\ 0 & 0 & t/z \end{bmatrix} \begin{bmatrix} (xa)/t & (xb)/t & (xc)/t \\ (ya)/t & (yb)/t & (yc)/t \\ (za)/t & (zb)/t & (zc)/t \end{bmatrix} \begin{bmatrix} 1/\sqrt{a/t} & 0 & 0 \\ 0 & 1/\sqrt{b/t} & 0 \\ 0 & 0 & 1/\sqrt{c/t} \end{bmatrix} \tag{60}$$

$$= \begin{bmatrix} (xa)/xt\sqrt{a/t} & (xb)/xt\sqrt{b/t} & (xc)/xt\sqrt{c/t} \\ (ya)/yt\sqrt{a/t} & (yb)/yt\sqrt{b/t} & (yc)/yt\sqrt{c/t} \\ (za)/zt\sqrt{a/t} & (zb)/zt\sqrt{b/t} & (zc)/zt\sqrt{c/t} \end{bmatrix} \tag{61}$$

よって

$$X = \left[\begin{array}{ccc} (xa)/xt\sqrt{a/t} & (xb)/xt\sqrt{b/t} & (xc)/xt\sqrt{c/t} \end{array} \right] \tag{62}$$

$$Y = \left[\begin{array}{ccc} (ya)/yt\sqrt{a/t} & (yb)/yt\sqrt{b/t} & (yc)/yt\sqrt{c/t} \end{array} \right] \tag{63}$$

$$Z = \left[\begin{array}{ccc} (za)/zt\sqrt{a/t} & (zb)/zt\sqrt{b/t} & (zc)/zt\sqrt{c/t} \end{array} \right] \tag{64}$$

さらに，これらの点から平均値を引くことで，平均 0 に中心化する．平均値は以下のように求められる．

たとえば点 ABC の 1 次元目 (x) について考える．散布図と同じように考えると，点 A には，A に該当する a 個の点が重なって布置されていると見なすことができる．BC も同様である．なので、点 A の 1 次元目を a_1 と置くと，平均値というのは，以下のように期待値の計算方法と同じように考えることができる．

$$\bar{x} = \frac{1}{t}(a \times a_1 + b \times b_1 + c \times c_1) \tag{65}$$

$$= \frac{a}{t}a_1 + \frac{b}{t}b_1 + \frac{c}{t}c_1 \tag{66}$$

ここで，$a_1 = (xa)/xt\sqrt{a/t}$ なので、,

$$\bar{x} = \frac{a}{t}\frac{(xa)}{a\sqrt{x/t}} + \frac{b}{t}\frac{(xb)}{b\sqrt{x/t}} + \frac{c}{t}\frac{(xc)}{c\sqrt{x/t}} \tag{67}$$

$$= \frac{(xa)}{t}\frac{\sqrt{t}}{\sqrt{x}} + \frac{(xb)}{t}\frac{\sqrt{t}}{\sqrt{x}} + \frac{(xc)}{t}\frac{\sqrt{t}}{\sqrt{x}} \tag{68}$$

$$= \frac{(xa) + (xb) + (xc)}{t}\frac{\sqrt{t}}{\sqrt{x}} \tag{69}$$

$$= \frac{x}{t}\frac{\sqrt{t}}{\sqrt{x}} = \sqrt{\frac{x}{t}} \tag{70}$$

となる．このように，平均値はその次元が表現する行もしくは列の周辺確率の平方根として求めることができる．よって，変換後の点はたとえば以下のようになる．

$$A = \left[\begin{array}{ccc} \frac{(xa)}{a\sqrt{x/t}} - \sqrt{\frac{x}{t}} & \frac{(ya)}{a\sqrt{y/t}} - \sqrt{\frac{y}{t}} & \frac{(za)}{a\sqrt{z/t}} - \sqrt{\frac{z}{t}} \end{array} \right] \tag{71}$$

$$\tag{72}$$

ここからは，これらの点を回転前の初期位置として用いる．

分散を最大にする回転法

具体的に，分散を最大にする手順を考えると，幾何学的には以下のようにするのが自然と言える．

- 条件を定めず適当に回転させて，分散が最大となる向きを見つけ，そこを 1 次元目とする．
- 1 次元目を固定した上で，その 1 次元目と直交する，分散が最大となる回転を見つけ，2 次元目とする．
- 以下，次元数-1 個の次元について，分散が大きいものから順になるように回転を行う．

こうすることで，分散が最大のものが 1 次元目，次が 2 次元目と定めることができる．次元数-1 個の回転軸が定まれば，最後の 1 個は確定するためそれ以上の回転を考える必要がない．具体的には，2 次元平面を回転させる場合，x 軸の向きが定まれば，それと直交する y 軸は自動的に定まっている．

目的関数の設定

点 A を以下のように置くとき，

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \tag{73}$$

3 つの点から分散を求める方法はいくつか考えられる．中でも直感的に最もわかりやすいのは，以下のように，点 A に該当するオブザベーション数 (a 回) だけ点 A の座標を繰り返し，次に b 回点 B の座標，同じように c 回点 C の座標を繰り返したデータ行列を作成する方法である．

$$\boldsymbol{X} = \begin{bmatrix} a_1 & a_2 & a_3 \\ & \vdots & \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ & \vdots & \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ & \vdots & \\ c_1 & c_2 & c_3 \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ & \vdots & \\ x_{t1} & x_{t2} & x_{t3} \end{bmatrix} \tag{74}$$

これは，平均値を求める方法と同じで，布置を散布図と見なすと，点 A には，a 個の点が重なっていると見なせることを用いている．こうすることで，1 行 1 オブザベーションのデータ行列として扱うことができ，主成分分析など，連続的なデータの分析と同じ考え方を用いることができる．

$$\bar{x}_1 = \bar{x}_2 = \bar{x}_3 = 0 \tag{75}$$

なので，分散は以下のように求められる．

$$V(\boldsymbol{X}) = \frac{1}{t}\boldsymbol{X}'\boldsymbol{X} \tag{76}$$

$$= \frac{1}{t}\begin{bmatrix} x_{11} & & x_{t1} \\ x_{12} & \cdots & x_{t2} \\ x_{13} & & x_{t3} \end{bmatrix}\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ & \vdots & \\ x_{t1} & x_{t2} & x_{t3} \end{bmatrix} \tag{77}$$

$$= \frac{1}{t}\begin{bmatrix} \sum_{i=1}^{t} x_{i1}^2 & \sum_{i=1}^{t} x_{i1}x_{i2} & \sum_{i=1}^{t} x_{i1}x_{i3} \\ \sum_{i=1}^{t} x_{i2}x_{i1} & \sum_{i=1}^{t} x_{i2}^2 & \sum_{i=1}^{t} x_{i2}x_{i3} \\ \sum_{i=1}^{t} x_{i3}x_{i1} & \sum_{i=1}^{t} x_{i3}x_{i2} & \sum_{i=1}^{t} x_{i3}^2 \end{bmatrix} \tag{78}$$

$$= \begin{bmatrix} \frac{1}{t}\sum_{i=1}^{t} x_{i1}^2 & \frac{1}{t}\sum_{i=1}^{t} x_{i1}x_{i2} & \frac{1}{t}\sum_{i=1}^{t} x_{i1}x_{i3} \\ \frac{1}{t}\sum_{i=1}^{t} x_{i2}x_{i1} & \frac{1}{t}\sum_{i=1}^{t} x_{i2}^2 & \frac{1}{t}\sum_{i=1}^{t} x_{i2}x_{i3} \\ \frac{1}{t}\sum_{i=1}^{t} x_{i3}x_{i1} & \frac{1}{t}\sum_{i=1}^{t} x_{i3}x_{i2} & \frac{1}{t}\sum_{i=1}^{t} x_{i3}^2 \end{bmatrix} \tag{79}$$

さて，ここで，座標の変換を考える．座標の変換は，行列 $\boldsymbol{W}$ を用いて以下のようになされる．

$$\boldsymbol{XW} = \boldsymbol{Y} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ & \vdots & \\ x_{t1} & x_{t2} & x_{t3} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} \tag{80}$$

$$= \begin{bmatrix} x_{11}w_{11} + x_{12}w_{21} + x_{13}w_{31} & x_{11}w_{12} + x_{12}w_{22} + x_{13}w_{32} & x_{11}w_{13} + x_{12}w_{23} + x_{13}w_{33} \\ & \vdots & \\ x_{t1}w_{11} + x_{t2}w_{21} + x_{t3}w_{31} & x_{t1}w_{12} + x_{t2}w_{22} + x_{t3}w_{32} & x_{t1}w_{13} + x_{t2}w_{23} + x_{t3}w_{33} \end{bmatrix} \tag{81}$$

1 番目の点の 1 次元目の座標は $x_{11}w_{11} + x_{12}w_{21} + x_{13}w_{31}$ と変換されている．つまり，1 次元目の変換にかかわる係数ベクトルは

$$\boldsymbol{w_1} = \begin{bmatrix} w_{11} \\ w_{21} \\ w_{31} \end{bmatrix} \tag{82}$$

ということになる．さて，ここで，変換の手順に戻ると，「条件を定めず適当に回転させて，分散が最大となる向きを見つけ，そこを 1 次元目とする.」である．つまり，変換

$$\boldsymbol{Xw_1} = \boldsymbol{y_1} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ & \vdots & \\ x_{t1} & x_{t2} & x_{t3} \end{bmatrix} \begin{bmatrix} w_{11} \\ w_{21} \\ w_{31} \end{bmatrix} \tag{83}$$

$$= \begin{bmatrix} x_{11}w_{11} + x_{12}w_{21} + x_{13}w_{31} \\ \vdots \\ x_{t1}w_{11} + x_{t2}w_{21} + x_{t3}w_{31} \end{bmatrix} \tag{84}$$

を行い，その分散 $V(y_1)$ を最大にする $\boldsymbol{w_1}$ を探すということになる．

$V(y_1)$ は，以下の通り平均値が 0 となるので，$V(\boldsymbol{X})$ と同じように求められる．

$$E(y_1) = \frac{1}{t}\sum_{i=1}^{t} x_{i1}w_{11} + x_{i2}w_{21} + x_{i3}w_{31} \tag{85}$$

$$= w_{11}\frac{1}{t}\sum_{i=1}^{t} x_{i1} + w_{21}\frac{1}{t}\sum_{i=1}^{t} x_{i2} + w_{31}\sum_{i=1}^{t} x_{i3} \tag{86}$$

$$= w_{11}E(x_1) + w_{21}E(x_2) + w_{31}E(x_3) = 0 + 0 + 0 = 0 \tag{87}$$

$$V(y_1) = \frac{1}{t}\boldsymbol{y_1}'\boldsymbol{y_1} = \frac{1}{t}(\boldsymbol{Xw_1})'\boldsymbol{Xw_1} = \frac{1}{t}\boldsymbol{w_1}'\boldsymbol{X}'\boldsymbol{Xw_1} = \boldsymbol{w_1}'V(\boldsymbol{X})\boldsymbol{w_1} \tag{88}$$

よって，分散を最大にするということは，$\boldsymbol{w_1}'V(\boldsymbol{X})\boldsymbol{w_1}$ を最大にする $\boldsymbol{w_1}$ を探す問題と考えることができる．ただし，$\boldsymbol{w_1}$ を大きくするとそれに伴い，$\boldsymbol{w_1}'V(\boldsymbol{X})\boldsymbol{w_1}$ はいくらでも大きくなってしまう．そこで，制約条件として，$\boldsymbol{w_1}$ を単位ベクトルとする．すなわちノルムが 1，

$$\sqrt{w_{11}^2 + w_{21}^2 + w_{31}^2} = w_{11}^2 + w_{21}^2 + w_{31}^2 = 1 \tag{89}$$

ということである．最大値を求めるには，ラグランジュの未定乗数法を用いるのがよい．

ラグランジュの未定乗数法

ラグランジュの未定乗数法は，関数 $f(\boldsymbol{x})$ を，$g(\boldsymbol{x}) = 0$ という条件の下で最大化したい場合に用いることができる手法である．

$$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x}) \tag{90}$$

を作り，$\boldsymbol{x}$ のサイズを n とするとき，$1 \leq i \leq n$ である任意の i について，以下を満たす点の中に $f(\boldsymbol{x})$ の極値が存在するというのがその定義である．

$$\frac{\partial L}{\partial x_i} = \frac{\partial L}{\partial \lambda} = 0 \tag{91}$$

尚，$g(\boldsymbol{x}) = 0, h(\boldsymbol{x}) = 0$ のように制約条件が 2 つ以上ある場合には，

$$L(\boldsymbol{x}, \lambda_1, \lambda_2) = f(\boldsymbol{x}) - \lambda_1 g(\boldsymbol{x}) - \lambda_2 h(\boldsymbol{x}) \tag{92}$$

として，

$$\frac{\partial L}{\partial x_i} = \frac{\partial L}{\partial \lambda_1} = \frac{\partial L}{\partial \lambda_2} = 0 \tag{93}$$

を満たす点を探す問題と考えればよい．

固有値問題への帰着

さて，回転におけるの分散の最大値算出にこの方法を用いる．データが与えられた下では $V(\boldsymbol{X})$ は既知であるので，最大値を求めたい関数は

$$f(\boldsymbol{w_1}) = \boldsymbol{w_1}' V(\boldsymbol{X}) \boldsymbol{w_1} \tag{94}$$

となる．また，制約条件は，$w_{11}^2 + w_{21}^2 + w_{31}^2 = 1$ を用いて

$$g(\boldsymbol{w_1}) = w_{11}^2 + w_{21}^2 + w_{31}^2 - 1 = 0 \tag{95}$$

と置ける．よって，

$$L(\boldsymbol{w_1}, \lambda_1) = \boldsymbol{w_1}' V(\boldsymbol{X}) \boldsymbol{w_1} - \lambda_1 (w_{11}^2 + w_{21}^2 + w_{31}^2 - 1) \tag{96}$$

について，

$$\frac{\partial L}{\partial w_{i1}} = \frac{\partial L}{\partial \lambda_1} = 0 \tag{97}$$

を満たす組み合わせを探すということになる．共分散行列は対称行列なので，

$$V(\boldsymbol{X}) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix} \tag{98}$$

と置くとき，

$$\boldsymbol{w_1}'V(\boldsymbol{X})\boldsymbol{w_1} = \begin{bmatrix} w_{11} & w_{21} & w_{31} \end{bmatrix} \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix} \begin{bmatrix} w_{11} \\ w_{21} \\ w_{31} \end{bmatrix} \tag{99}$$

$$= \begin{bmatrix} w_{11}\sigma_1^2 + w_{21}\sigma_{12} + w_{31}\sigma_{13} & w_{11}\sigma_{12} + w_{21}\sigma_2^2 + w_{31}\sigma_{23} & w_{11}\sigma_{13} + w_{21}\sigma_{23} + w_{31}\sigma_3^2 \end{bmatrix} \begin{bmatrix} w_{11} \\ w_{21} \\ w_{31} \end{bmatrix} \tag{100}$$

$$= w_{11}(w_{11}\sigma_1^2 + w_{21}\sigma_{12} + w_{31}\sigma_{13}) + w_{21}(w_{11}\sigma_{12} + w_{21}\sigma_2^2 + w_{31}\sigma_{23}) + w_{31}(w_{11}\sigma_{13} + w_{21}\sigma_{23} + w_{31}\sigma_3^2) \tag{101}$$

$$= w_{11}^2\sigma_1^2 + w_{11}w_{21}\sigma_{12} + w_{11}w_{31}\sigma_{13} + w_{11}w_{21}\sigma_{12} + w_{21}^2\sigma_2^2 + w_{21}w_{31}\sigma_{23} + w_{11}w_{31}\sigma_{13} + w_{21}w_{31}\sigma_{23} + w_{31}^2\sigma_3^2 \tag{102}$$

$$= w_{11}^2\sigma_1^2 + w_{21}^2\sigma_2^2 + w_{31}^2\sigma_3^2 + 2w_{11}w_{21}\sigma_{12} + 2w_{11}w_{31}\sigma_{13} + 2w_{21}w_{31}\sigma_{23} \tag{103}$$

なので，

$$L(\boldsymbol{w_1}, \lambda_1) = w_{11}^2\sigma_1^2 + w_{21}^2\sigma_2^2 + w_{31}^2\sigma_3^2 + 2w_{11}w_{21}\sigma_{12} + 2w_{11}w_{31}\sigma_{13} + 2w_{21}w_{31}\sigma_{23} - \lambda_1 w_{11}^2 - \lambda_1 w_{21}^2 - \lambda_1 w_{31}^2 + \lambda_1 \tag{104}$$

となり，以下の 4 つの連立方程式が成立する．

$$\frac{\partial L}{\partial w_{11}} = 2w_{11}\sigma_1^2 + 2w_{21}\sigma_{12} + 2w_{31}\sigma_{13} - 2\lambda_1 w_{11} = 0 \tag{105}$$

$$\frac{\partial L}{\partial w_{21}} = 2w_{11}\sigma_{12} + 2w_{21}\sigma_2^2 + 2w_{31}\sigma_{23} - 2\lambda_1 w_{21} = 0 \tag{106}$$

$$\frac{\partial L}{\partial w_{31}} = 2w_{11}\sigma_{13} + 2w_{21}\sigma_{23} + 2w_{31}\sigma_3^2 - 2\lambda_1 w_{31} = 0 \tag{107}$$

$$\frac{\partial L}{\partial \lambda_1} = 1 - w_{11}^2 - w_{21}^2 - w_{31}^2 = 0 \tag{108}$$

λ_1 に関する偏微分を除く 3 つをまとめると，

$$\frac{\partial L}{\partial \boldsymbol{w_1}} = 2\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix} \begin{bmatrix} w_{11} \\ w_{21} \\ w_{31} \end{bmatrix} - 2\lambda_1 \begin{bmatrix} w_{11} \\ w_{21} \\ w_{31} \end{bmatrix} \tag{109}$$

$$= 2V(\boldsymbol{X})\boldsymbol{w_1} - 2\lambda_1\boldsymbol{w_1} = \boldsymbol{0} \tag{110}$$

$$V(\boldsymbol{X})\boldsymbol{w_1} = \lambda_1\boldsymbol{w_1} \tag{111}$$

が得られる．$1 - w_{11}^2 - w_{21}^2 - w_{31}^2 = 0$ であることも考慮すると，これは，共分散行列 $V(\boldsymbol{X})$ を固有値分解したときに得られる固有ベクトルを，$1 - w_{11}^2 - w_{21}^2 - w_{31}^2 = 0$ を満たすように調整して得られた結果が係数ベクトル $\boldsymbol{w_1}$ になることを意味している．また，$V(\boldsymbol{X})\boldsymbol{w_1} = \lambda_1\boldsymbol{w_1}$ の両辺に左から $\boldsymbol{w_1}'$ をかけると，

$$\boldsymbol{w_1}'V(\boldsymbol{X})\boldsymbol{w_1} = \boldsymbol{w_1}'\lambda_1\boldsymbol{w_1} \tag{112}$$

$$V(\boldsymbol{y_1}) = \begin{bmatrix} w_{11} & w_{21} & w_{31} \end{bmatrix} \lambda_1 \begin{bmatrix} w_{11} \\ w_{21} \\ w_{31} \end{bmatrix} \tag{113}$$

$$= \lambda_1(w_{11}^2 + w_{21}^2 + w_{31}^2) \tag{114}$$

であり，$w_{11}^2 + w_{21}^2 + w_{31}^2 = 1$ なので，

$$V(\boldsymbol{y_1}) = \lambda_1 \tag{115}$$

である．よって，変換して得られた新しい変数の分散は固有値 λ_1 と一致する．ここから，$w_{11}^2 + w_{21}^2 + w_{31}^2 = 1$ という条件で共分散行列を固有値分解して得られた固有値のうち，最大となる値が変換した後の 1 次元目の分散になる．また，その固有値に対応する固有ベクトルが 1 次元目を構成する変換のための係数ということになる．

2 次元目以降の導出

では，2 番目以降の大きさの固有値の意味はどうなるであろうか．それを考えるために，変換により 2 次元目を求める．2 次元目の係数ベクトルは以下であった．

$$\boldsymbol{w_2} = \begin{bmatrix} w_{12} \\ w_{22} \\ w_{32} \end{bmatrix} \tag{116}$$

元のデータ行列に，この係数ベクトルをかけた結果が変換後の 2 次元目の座標になるのは，1 次元目と同じである．よって，その分散も同じように求められ，また，制約条件も同じように設定できる．それぞれ以下の通りである．

$$\boldsymbol{X}\boldsymbol{w_2} = \boldsymbol{y_2} \tag{117}$$

$$V(y_2) = \boldsymbol{w_2}'V(\boldsymbol{X})\boldsymbol{w_2} \tag{118}$$

$$0 = 1 - w_{12}^2 - w_{22}^2 - w_{32}^2 \tag{119}$$

ただし，1 次元目は 2 次元目と直交している必要がある．これは，係数ベクトル同士の内積が 0 であればよいので，

$$\boldsymbol{w_1}'\boldsymbol{w_2} = w_{11}w_{12} + w_{21}w_{22} + w_{31}w_{32} = 0 \tag{120}$$

という制約条件が新たに加わることを意味する．この条件を付け加えてラグランジュの未定乗数法を考える．

$$L(\boldsymbol{w_2}, \lambda_2, \mu) = \boldsymbol{w_2}'V(\boldsymbol{X})\boldsymbol{w_2} - \lambda_2(w_{12}^2 + w_{22}^2 + w_{32}^2 - 1) - \mu(w_{11}w_{12} + w_{21}w_{22} + w_{31}w_{32}) \tag{121}$$

これも $\boldsymbol{w_2}$ について偏微分する．右辺最初の 2 項は 1 次元目と基本的に変わっていないので，

$$\frac{\partial L}{\partial \boldsymbol{w_2}} = 2V(\boldsymbol{X})\boldsymbol{w_2} - 2\lambda_2\boldsymbol{w_2} - \frac{\partial}{\partial \boldsymbol{w_2}}\mu(w_{11}w_{12} + w_{21}w_{22} + w_{31}w_{32}) \tag{122}$$

$$= 2V(\boldsymbol{X})\boldsymbol{w_2} - 2\lambda_2\boldsymbol{w_2} - \mu\boldsymbol{w_1} = \boldsymbol{0} \tag{123}$$

$$\tag{124}$$

である．さてここで，この式に左から $\boldsymbol{w_1}'$ をかけることを考える．

$$2\boldsymbol{w_1}'V(\boldsymbol{X})\boldsymbol{w_2} - 2\lambda_2\boldsymbol{w_1}'\boldsymbol{w_2} - \mu\boldsymbol{w_1}'\boldsymbol{w_1} = \boldsymbol{w_1}'\boldsymbol{0} \tag{125}$$

$$2\boldsymbol{w_1}'V(\boldsymbol{X})\boldsymbol{w_2} - 2\lambda_2 \times 0 - \mu \times 1 = 0 \tag{126}$$

$$\mu = 2\boldsymbol{w_1}'V(\boldsymbol{X})\boldsymbol{w_2} \tag{127}$$

ここで $V(\boldsymbol{X})$ は対称行列なので，

$$\mu = 2\boldsymbol{w_2}'V(\boldsymbol{X})\boldsymbol{w_1} \tag{128}$$

と置き換えられる．すると，$V(\boldsymbol{X})\boldsymbol{w_1} = \lambda_1\boldsymbol{w_1}$ なので，

$$\mu = 2\boldsymbol{w_2}'\lambda_1\boldsymbol{w_1} = 2\lambda_1 \times 0 = 0 \tag{129}$$

が導かれる．よって，

$$2V(\boldsymbol{X})\boldsymbol{w_2} - 2\lambda_2\boldsymbol{w_2} - \mu\boldsymbol{w_1} = \boldsymbol{0} \tag{130}$$

$$2V(\boldsymbol{X})\boldsymbol{w_2} - 2\lambda_2\boldsymbol{w_2} - \boldsymbol{0} = \boldsymbol{0} \tag{131}$$

$$V(\boldsymbol{X})\boldsymbol{w_2} = \lambda_2\boldsymbol{w_2} \tag{132}$$

となり，これもまた固有値問題に帰結する．よって，固有値の中で 2 番目に大きな値を λ_2 とし，それが 2 次元目の分散となる．それに対応する固有ベクトル $\boldsymbol{w_2}$ は 2 次元目の座標を求めるための係数ベクトルである．以下，3 次元目以上の次元 n についても，$n-1$ 次元目までの係数ベクトルとの内積を 0 とするという制約条件を追加してラグランジュの未定乗数法を用い，その式を整理することですべて固有値問題の形にすることができる．よって，最終的には以下の流れで係数行列 $\boldsymbol{W}$ を求めることができる．

- 回転前の行列から共分散行列を求める．
- 共分散行列を固有値分解する．
- 得られた固有値を大きなものから順に並べる．この時の固有値が各次元の分散である．
- 固有値に対応する固有ベクトルがそれぞれの次元の係数ベクトルとなる．それらを並べて係数行列を構成する．i 番目に大きな値の固有値に対応する固有ベクトルを $\boldsymbol{w_i}$ としたとき，係数行列は以下の通りである．

$$\boldsymbol{W} = \begin{bmatrix} w_{11} & \cdots & w_{1i} & \cdots & w_{1n} \\ \vdots & & \vdots & & \vdots \\ w_{i1} & \cdots & w_{ii} & \cdots & w_{in} \\ \vdots & & \vdots & & \vdots \\ w_{n1} & \cdots & w_{ni} & \cdots & w_{nn} \end{bmatrix} = \begin{bmatrix} \boldsymbol{w_1} & \cdots & \boldsymbol{w_i} & \cdots & \boldsymbol{w_n} \end{bmatrix} \tag{133}$$

実データの分析

本節では，テレビに重視する機能の各カテゴリについて，最適な布置を考える．座標の再定義に用いた方法で表 1 を変換すると，列，ジャンル側の初期座標は表 5 のようになる．この座標について，2 次元目までを用いて 2 次元平面上に布置すると図 17 の通りである．

表 5：機能における各カテゴリの初期座標

	1 次元目	2 次元目	3 次元目
価格	-0.0492	-0.0508	-0.0295
画面サイズ	-0.0194	-0.0146	0.0536
録画機能	0.0648	-0.0302	-0.0301

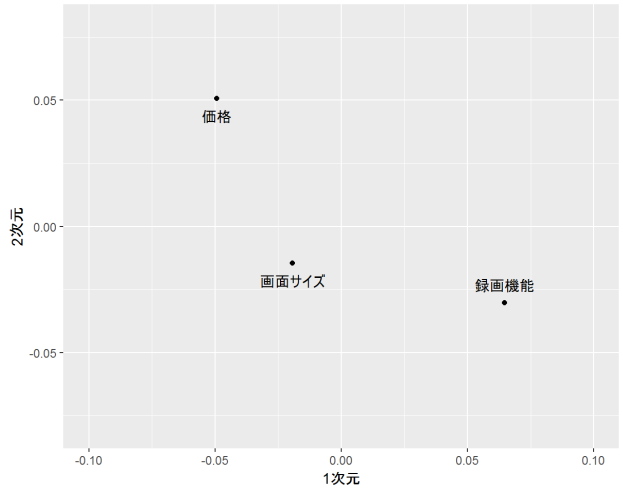

図 17：回転の数値例

ここから共分散行列を求めると，

$$V(\boldsymbol{X}) = \begin{bmatrix} 0.0023 & -0.0013 & -0.0006 \\ -0.0013 & 0.0017 & -0.0004 \\ -0.0006 & -0.0004 & 0.0016 \end{bmatrix} \tag{134}$$

となる．よって，この行列の固有値と固有ベクトルを求めればよい．固有値はそれぞれ 0.0032，0.0018，0.0000 であり，これが各次元の分散の大きさとなる．対応する固有ベクトルを横に並べた係数行列は以下の通りである．

$$\boldsymbol{W} = \begin{bmatrix} 0.8429 & -0.0716 & -0.5332 \\ -0.5091 & -0.4266 & -0.7475 \\ -0.1740 & 0.9016 & -0.3961 \end{bmatrix} \tag{135}$$

この変換前の行列にこの行列をかけることで，変換後の座標を求めることができる．即ち．

$$\boldsymbol{XW} = \begin{bmatrix} -0.0492 & -0.0508 & -0.0295 \\ -0.0194 & -0.0146 & 0.0536 \\ 0.0648 & -0.0302 & -0.0301 \end{bmatrix} \begin{bmatrix} 0.8429 & -0.0716 & -0.5332 \\ -0.5091 & -0.4266 & -0.7475 \\ -0.1740 & 0.9016 & -0.3961 \end{bmatrix} = \begin{bmatrix} -0.0622 & -0.0448 & 0 \\ -0.0183 & 0.0559 & 0 \\ 0.0752 & -0.0189 & 0 \end{bmatrix} \tag{136}$$

である．よって，この 3 点を 2 次元平面上に布置すると図 18 となる．

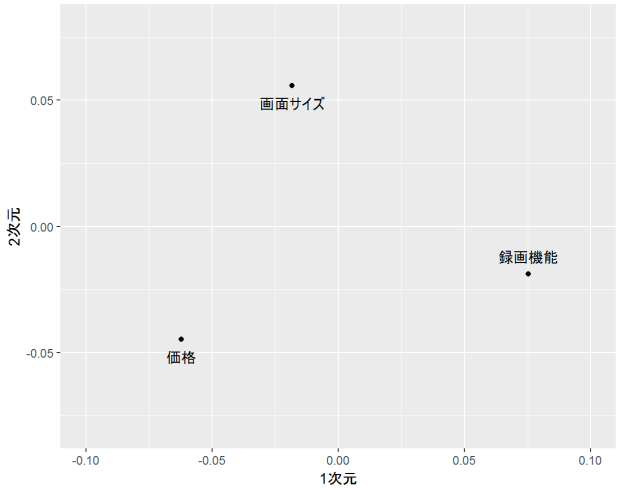

図 18：回転後の布置

この結果は，R のパッケージを用いた分析結果である図 6 と，上下左右がそれぞれ入れ替わったものになっている．固有値分解では，固有ベクトルの方向が 180 度変わっても同じ結果となるため，こういったことは珍しくない．

まとめ

コレスポンデンス分析では，分割表からカテゴリ間の χ^2 距離を求めることで，カテゴリ間の関係を把握する．そして，その距離の情報をなるべく維持した形で低次元空間にその情報を縮約，布置するという方法をとっている．そのために，χ^2 距離の情報を持った座標から各次元の共分散行列を求め，それを固有値分解し，固有ベクトルから変換行列を作成，元の行列にかけて変換を行っている．

ただし，この方法では，分析する度に上下左右がそれぞれ入れ替わる可能性がある．これは，固有値分解の性質であり，分析の際には注意する必要がある．

文献

Clausen,S.E.(藤本一男訳)(2015).Applied Correspondence Analysis An Introduction(対応分析入門 原理から応用まで)，オーム社.

樋口耕一 (2004). テキスト型データの計量的分析：2 つのアプローチの峻別と統合，理論と方法 19(1), 101-115.