

第一主成分を選んでいるのは、データではなく単位である

— 同じ標本を cm で測れば **55.6503** 度、m なら **89.7345** 度、mm なら **4.1934** 度を向く — 数値は一つも変えていない。掛けただけである — 標準化すれば **45.0000** 度に落ち着くが、標準化もまた一つの選択である — [論文375]

山岸 勇輝 (Yuuki YAMAGISHI)

Yamagishi Complexity Torus Lab, Japan | 山岸勇輝ラボ

2026年9月

— From thinking to syncing. —

要旨

主成分分析は「データの中の主要な方向」を取り出すと説明される。本稿が問うのはその方向がデータだけで決まるかである——答えは決まらない。単位が決めている。新しい定理も法則も主張しない。第一に、身長と体重の標本を作り、単位だけを変えて主成分分析にかける（第2節）。第二に、これが本稿の芯である。cm で **55.6503** 度、m で **89.7345** 度、mm で **4.1934** 度を向く（第2節）。第三に、m と mm では、第一主成分がほぼ直交する二つの軸を指す（第2節）。第四に、寄与率も動き、m で **1.0000**、mm で **0.9898** になる（第3節）。第五に、標準化すれば **45.0000** 度・寄与率 **0.7899** に落ち着く（第4節）。第六に、分離子は「共分散行列か相関行列か」であって、データではない（第5節）。

本稿の射程（射程注記）：新しい定理も法則も主張しない——主成分分析・共分散行列と相関行列の違い・標準化はすべて既知である。多変量解析を作らない——主成分の個数の決め方や、因子分析との関係には立ち入らない。名前を挙げるにとどめる。実データを扱わない——実際の身体測定は用いない。相関 **0.6** の二変量正規から生成した標本を使う。推定論を作らない——固有値の標本誤差や検定は主題ではない。三変数以上を扱わない——二変数の場合だけを計算する。変数が増えても構造は同じである。解釈を与えない——第一主成分に「体格」などの名前を付けることはしない。既刊との関係：論文358 は被写界深度の数が約束で決まることを示した——そこでは約束が許容錯乱円だった。ここでは単位である。論文356 は次元が同じでも別の量であることを示した——ここでは次元の違う量を同じ空間に並べている。論文300 は同根か別根かが判定できることを示した——「データの構造」と「単位の選択」は、その基準で別根である。論文372 は指数が結論を支配することを示した——ここでは尺度が結論を支配する。加えたのは、四つの尺度で第一主成分の向きと寄与率を数に出したこと、m と mm でほぼ直交することを示したこと、標準化すると **45** 度に落ち着くことを示したこと、寄与率が **0.7899** から **1.0000** まで動くことを示したこと、分離子を「共分散行列か相関行列か」に置いたことである。

1. 記法と問い

設定：身長（平均 170 cm・標準偏差 8 cm）と体重（平均 65 kg・標準偏差 10 kg）を、相関 0.6 で 4000 個生成する。

$$\Sigma = \begin{pmatrix} \sigma_h^2 & \rho\sigma_h\sigma_w \\ \rho\sigma_h\sigma_w & \sigma_w^2 \end{pmatrix}, \quad \text{第一主成分} = \Sigma \text{ の最大固有値に属する固有ベクトル}$$

観察 1.1

主成分分析は分散の大きい方向を探す。共分散行列の固有ベクトル——「大きい」は数値の大きさに測られる。

観察 1.2

だが数値の大きさは単位で変わる。170 cm は 1.70 m でもあり 1700 mm でもある——同じ長さが、 10^6 倍の分散の違いを生む。

観察 1.3

そして体重の単位は変えない。kg のまま——二つの軸の「大きさ」の比が動く。本稿はその帰結を数で出す。

問い：第一主成分の向きは、データだけで決まるか。

2. cm・m・mm で、向きが 85 度動く

計算 2.1：単位に対する第一主成分（本稿で計算・同じ標本 4000 個）。

単位	第一主成分の向き [度]	寄与率	第一固有値
cm と kg	55.6503	0.8035	129.454026
m と kg	89.7345	1.0000	98.318922
mm と kg	4.1934	0.9898	6312.309583

観察 2.2

これが本稿の芯である。m で 89.7345 度、mm で 4.1934 度——差は 85.5411 度で、ほぼ直交する。同じ標本、同じ人、同じ測定である。

観察 2.3

向きの意味も入れ替わる。89.7345 度はほぼ体重の軸、4.1934 度はほぼ身長 of 軸——「主要な変動は体重だ」と「身長だ」が、単位 of 選択だけで入れ替わる。

観察 2.4

理由は単純である。mm にすると身長 of 分散が 10^6 倍になる——体重 of 分散を圧倒するので、第一主成分が身長 of 軸に張り付く。

観察 2.5

m にすると逆になる。身長 of 分散が 10^{-4} 倍になる——今度は体重が圧倒する。

観察 2.6

cm はその中間である。55.6503 度——「ちょうどよい」ように見えるが、cm を選ぶ理由はデータの側には無い。

3. 寄与率も動く

観察 3.1

寄与率は m で 1.0000、mm で 0.9898 である。どちらも「第一主成分でほぼ全部説明できる」と読める——だが説明している軸が別である。

観察 3.2

これがいちばん誤解を招く。寄与率が高いことは、その方向が意味をもつことを保証しない——単位を極端にすれば、寄与率はいくらでも 1 に近づく。

観察 3.3

第一固有値そのものも比較できない。**98.318922** と **6312.309583**——単位が違えば固有値の単位も違うので、大小に意味が無い。

観察 3.4

cm の寄与率 **0.8035** が、いちばん低い。三つの中で——「説明力が低い」のではなく、二つの軸の分散が近いというだけである。

4. 標準化すると落ち着く

計算 4.1：両方の変数を標準化した場合（本稿で計算）。

処理	向き [度]	寄与率	第一固有値
標準化（両方）	45.0000	0.7899	1.579873

観察 4.2

ちょうど **45.0000** 度になる。二つの分散が等しくなるからである——相関行列の固有ベクトルは、二変数なら常に **45** 度と **135** 度である。

観察 4.3

寄与率 **0.7899** は相関に対応する。相関 **0.6** なら $(1 + \rho)/2 = 0.8$ ——標本の相関がやや低いぶん **0.7899** になっている。式が閉じている。

観察 4.4

標準化すれば単位に依存しなくなる。cm でも m でも mm でも同じ結果——これが相関行列を使う理由である。

観察 4.5

正直に言えば——標準化もまた一つの選択である。「全ての変数を同じ重みにする」という決定であって、中立ではない——分散の大きい変数を重く見たい場合には、むしろ共分散行列が正しい。

観察 4.6

つまり選択は消えず、場所を変える。単位を選ぶか、標準化するかを選ぶか——論文358 で許容錯乱円の約束が場所を変えたただけだったのと同じ形である。

5. 分離子：共分散行列か、相関行列か

観察 5.1

分離子——共分散行列を使うか相関行列を使うかであって、データではない。

	単位に依存するか	二変数での向き	何を重く見るか
共分散行列	する	4.19～89.73 度	分散の大きい変数
相関行列	しない	常に 45 度	全変数を等しく

観察 5.2

右端の列が、どちらを使うべきかを定める。単位に意味がある場合（同じ単位の複数の測定）は共分散行列——単位がばらばらなら相関行列。「どちらが正しいか」ではなく「何を重く見たいか」である。

観察 5.3

論文300の基準で別根である。データの構造は標本から、尺度の選択は分析者から来る——結果はその積であって、片方だけの性質ではない。

観察 5.4

分離子は動かせる。標準化すれば単位依存が消える——ただし「等しい重み」という新しい仮定が入る。消えたのは依存であって、選択ではない。

観察 5.5

正直に言えば——55.6503度も0.7899も、相関0.6・4000標本という明示した条件の上での値である。別の標本なら別の数が出る——動かないのは、共分散行列の主成分が単位で動き、相関行列のそれが動かないという構造だけである。

観察 5.6

最後に一つ——「データが語る」と言うとき、語らせている尺度は誰かが選んでいる。単位を変えただけで第一主成分が85度回るのだから、その方向はデータだけの性質ではない——分析の結果を報告するときは、尺度も一緒に報告しなければ、再現できない。

6. 点呼と結び

1. cmで55.6503度（計算2.1）：中間。
2. mで89.7345度（計算2.1）：ほぼ体重の軸。
3. mmで4.1934度（計算2.1）：ほぼ身長 of 軸。
4. 差は85.5411度（観察2.2）：ほぼ直交する。
5. 数値は掛けただけ（観察2.2）：同じ標本・同じ人。
6. 寄与率はmで1.0000（観察3.1）：mmで0.9898。
7. 高い寄与率は意味を保証しない（観察3.2）：単位で作れる。
8. 標準化で45.0000度（計算4.1）：二変数なら常に。

9. 寄与率 **0.7899** は $(1 + \rho)/2$ に対応（観察 4.3）：式が閉じている。

10. 標準化も一つの選択（観察 4.5）：中立ではない。

結び：主成分分析は「データの中の主要な方向」を取り出すと説明される——だがその方向は、データだけでは決まらない。主成分分析は分散の大きい方向を探し、「大きい」は数値の大きさに測られる——ところが数値の大きさは単位で変わる。**170 cm** は **1.70 m** でもあり **1700 mm** でもあって、同じ長さが 10^6 倍の分散の違いを生む。だから同じ標本で、向きが **85 度** 動く——体重を **kg** に固定したまま身長単位の単位だけを変えると、第一主成分は **cm** で **55.6503 度**、**m** で **89.7345 度**、**mm** で **4.1934 度** を向く。**m** と **mm** の差は **85.5411 度** で、ほぼ直交する。**89.7345 度** はほぼ体重の軸、**4.1934 度** はほぼ身長の軸なので、「主要な変動は体重だ」と「身長だ」が単位の選択だけで入れ替わる。同じ標本、同じ人、同じ測定であり、数値は掛けただけである。寄与率も動く——**m** で **1.0000**、**mm** で **0.9898**。どちらも「第一主成分でほぼ全部説明できる」と読めるが、説明している軸が別である。高い寄与率は、その方向が意味をもつことを保証しない——単位を極端にすればいくらかでも **1** に近づく。第一固有値そのもの（**98.318922** と **6312.309583**）も、単位が違えば単位が違うので比較できない。標準化すれば落ち着く——両方を標準化すると向きはちょうど **45.0000 度**、寄与率は **0.7899** になる。二つの分散が等しくなるので、相関行列の固有ベクトルは二変数なら常に **45 度** である。寄与率も $(1 + \rho)/2$ に対応しており、式が閉じている。分けているものは一つ——共分散行列を使うか相関行列を使うかであって、データではない。単位に意味がある場合は共分散行列、単位がばらばらなら相関行列で、「どちらが正しいか」ではなく「何を重く見たいか」である。論文300の基準で別根であって、データの構造は標本から、尺度の選択は分析者から来る。結果はその積である。正直に言えば——標準化もまた一つの選択である。「全ての変数を同じ重みにする」という決定であって中立ではなく、分散の大きい変数を重く見たい場合にはむしろ共分散行列が正しい。選択は消えず、場所を変えるだけである。最後に一つ——「データが語る」と言うとき、語らせている尺度は誰かが選んでいる。単位を変えただけで第一主成分が **85 度** 回るのだから、その方向はデータだけの性質ではない。分析の結果を報告するときは、尺度も一緒に報告しなければ、再現できない。

参考文献

1. Pearson, K. (1901). On Lines and Planes of Closest Fit to Systems of Points in Space. *Philosophical Magazine*.
2. Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*.
3. Jolliffe, I. T. (2002). *Principal Component Analysis* (2nd ed.). New York: Springer.
4. Jolliffe, I. T., Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A*.
5. Mardia, K. V., Kent, J. T., Bibby, J. M. (1979). *Multivariate Analysis*. London: Academic Press.
6. 山岸勇輝 (2026). 『第三百回の点呼——同根か別根かは判定できる』[論文300]. Zenodo. 10.5281/zenodo.22171914
7. 山岸勇輝 (2026). 『電流は流れているのに、仕事をしていない』[論文356]. Zenodo. 10.5281/zenodo.22697124
8. 山岸勇輝 (2026). 『被写界深度の「深さ」は、測定値ではなく約束である』[論文358]. Zenodo. 10.5281/zenodo.22697216

9. 山岸勇輝 (2026). 『面積を半分にしても、種は半分にならない——そして種を半分にするには面積を 6.25 パーセントにする』 [論文372]. Zenodo. 10.5281/zenodo.22698241

— *From thinking to syncing.* —

山岸 勇輝 (Yuuki YAMAGISHI)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)