

What Chooses the First Principal Component Is Not the Data but the Units

— The same sample points at **55.6503** degrees in cm, **89.7345** in m, and **4.1934** in mm — Not one value was changed; they were multiplied — Standardising settles it at **45.0000** degrees, but standardising is itself a choice

— [Paper 375]

Yuuki YAMAGISHI (山岸 勇輝)

Yamagishi Complexity Torus Lab, Japan

September 2026

— *From thinking to syncing.* —

Abstract

Principal component analysis is described as extracting the main direction in the data. This paper asks whether the data alone fix that direction—the answer is that they do not: the units do. No new mathematical theorem and no new law is claimed. First, a sample of heights and weights is generated and analysed with only the units changed (Section 2). Second, this is the core of the paper. It points at **55.6503** degrees in cm, **89.7345** in m, and **4.1934** in mm (Section 2). Third, in m and in mm the first component picks out two nearly orthogonal axes (Section 2). Fourth, the explained fraction also moves, reaching 1.0000 in m and 0.9898 in mm (Section 3). Fifth, standardising settles it at **45.0000** degrees with a fraction of **0.7899** (Section 4). Sixth, the separator is covariance matrix against correlation matrix, not the data (Section 5).

Scope of this paper (scope note): No new mathematical theorem and no new law is claimed——principal component analysis, the difference between covariance and correlation matrices, and standardisation are all standard. We do not build multivariate analysis——how many components to retain, and the relation to factor analysis, are not entered. They are named and no more. We use no real data——actual body measurements are not used. A sample from a bivariate normal with correlation 0.6 is generated. We build no estimation theory——sampling error in eigenvalues and tests on them are not the subject. We do not treat more than two variables——only the bivariate case is computed. The structure is the same with more. We offer no interpretation——the first component is not given a name such as “build”. Relation to earlier papers: Paper 358 showed a depth-of-field number fixed by a convention——there the convention was the circle of confusion; here it is the unit. Paper 356 showed quantities of the same dimension to be different quantities——here quantities of different dimensions are put in one space. Paper 300 showed whether two things share a root is decidable——by that test, “the structure of the data” and “the choice of scale” have distinct roots. Paper 372 showed an exponent governing the conclusion——here the scale governs it. What is added is giving the direction and explained fraction under four scalings, showing m and mm to be nearly orthogonal, showing standardisation settling at 45 degrees, showing the explained fraction running from 0.7899 to 1.0000, and putting the separator on covariance against correlation.

1. Notation and the Question

Setting: Generate 4000 pairs of height (mean 170 cm, s.d. 8 cm) and weight (mean 65 kg, s.d. 10 kg) with correlation 0.6.

$$\Sigma = \begin{pmatrix} \sigma_h^2 & \rho\sigma_h\sigma_w \\ \rho\sigma_h\sigma_w & \sigma_w^2 \end{pmatrix}, \quad \text{first component} = \text{eigenvector of the largest eigenvalue of } \Sigma$$

Observation 1.1

PCA looks for the direction of greatest variance. An eigenvector of the covariance matrix——and “greatest” is measured by the size of the numbers.

Observation 1.2

But the size of numbers changes with units. 170 cm is also 1.70 m and 1700 mm——one length producing variances that differ by 10^6 .

Observation 1.3

And the weight's unit is left alone. Still kg——so the ratio of the two axes' “sizes” moves. This paper gives the consequence as numbers.

Question: Do the data alone fix the direction of the first component?

2. In cm, m and mm, the Direction Moves 85 Degrees

Calculation 2.1: First component against units (computed in this paper, one sample of 4000).

Units	Direction [deg]	Explained fraction	First eigenvalue
cm and kg	55.6503	0.8035	129.454026
m and kg	89.7345	1.0000	98.318922
mm and kg	4.1934	0.9898	6312.309583

Observation 2.2

This is the core of the paper. 89.7345 degrees in m and 4.1934 in mm——a gap of 85.5411 degrees, nearly orthogonal. The same sample, the same people, the same measurements.

Observation 2.3

The meaning swaps too. 89.7345 degrees is nearly the weight axis; 4.1934 nearly the height axis——“the main variation is weight” and “it is height” swap on the choice of unit alone.

Observation 2.4

The reason is simple. In mm the height variance is 10^6 times larger——it overwhelms the weight variance, so the first component sticks to the height axis.

Observation 2.5

In m it reverses. The height variance becomes 10^{-4} times—now the weight overwhelms.

Observation 2.6

Centimetres sit in between. 55.6503 degrees—it looks “just right”, but nothing on the data’s side gives a reason to choose cm.

3. The Explained Fraction Moves Too

Observation 3.1

It is 1.0000 in m and 0.9898 in mm. Both read as “the first component explains nearly everything”—and the axis being explained is different.

Observation 3.2

This is the most misleading part. A high explained fraction does not guarantee that the direction means anything—push the units far enough and it approaches 1 as closely as you like.

Observation 3.3

The eigenvalues themselves are not comparable either. 98.318922 against 6312.309583—different units make the eigenvalue’s units different, so their sizes mean nothing across rows.

Observation 3.4

The cm fraction, 0.8035, is the lowest of the three. Not because it explains less—but because the two axes’ variances are closest there.

4. Standardising Settles It

Calculation 4.1: With both variables standardised (computed in this paper).

Treatment	Direction [deg]	Explained fraction	First eigenvalue
Standardised (both)	45.0000	0.7899	1.579873

Observation 4.2

It is exactly 45.0000 degrees. Because the two variances become equal—the eigenvectors of a correlation matrix are always at 45 and 135 degrees in two variables.

Observation 4.3

The fraction 0.7899 corresponds to the correlation. For $\rho = 0.6$, $(1 + \rho)/2 = 0.8$ —the sample correlation is slightly lower, giving 0.7899. The formula closes.

Observation 4.4

Standardised, it no longer depends on units. The same result in cm, m and mm—this is why the correlation matrix is used.

Observation 4.5

To be stated honestly—standardising is itself a choice. It is the decision to give every variable equal weight, which is not neutral—where a high-variance variable ought to weigh more, the covariance matrix is the right one.

Observation 4.6

So the choice does not vanish; it changes place. Choose the units, or choose whether to standardise—the same shape as the circle-of-confusion convention merely changing place in Paper 358.

5. The Separator: Covariance Matrix, or Correlation Matrix?

Observation 5.1

The separator——whether the covariance or the correlation matrix is used; not the data.

	Unit-dependent	Direction in two variables	What is weighted
Covariance matrix	yes	4.19 to 89.73 deg	high-variance variables
Correlation matrix	no	always 45 deg	all variables equally

Observation 5.2

The rightmost column decides which to use. Where units are meaningful (several measurements in the same unit), the covariance matrix——where units are mixed, the correlation matrix. Not “which is right” but “what should weigh more”.

Observation 5.3

Distinct roots by the criterion of Paper 300. The structure of the data comes from the sample; the choice of scale from the analyst——the result is their product, not a property of either alone.

Observation 5.4

The separator can be moved. Standardising removes the unit dependence——but it introduces the new assumption of equal weight. What vanished is the dependence, not the choice.

Observation 5.5

To be stated honestly——55.6503 degrees and 0.7899 are values under stated conditions: correlation 0.6 and 4000 draws. Another sample gives other numbers——what does not move is the structure: covariance components move with units and correlation components do not.

Observation 5.6

One last thing——when one says “the data speak”, someone chose the scale in which they speak. Changing a unit alone turns the first component by 85 degrees, so that direction is not a property of the data alone——report an analysis without its scale and it cannot be reproduced.

6. The Census and the Close

1. **55.6503 degrees in cm** (Calculation 2.1): **in between**.
2. **89.7345 degrees in m** (Calculation 2.1): **nearly the weight axis**.
3. **4.1934 degrees in mm** (Calculation 2.1): **nearly the height axis**.
4. A gap of **85.5411 degrees** (Observation 2.2): **nearly orthogonal**.
5. The values were only **multiplied** (Observation 2.2): **same sample, same people**.
6. Fraction **1.0000 in m** (Observation 3.1): **0.9898 in mm**.
7. A **high fraction guarantees nothing** (Observation 3.2): **units can manufacture it**.
8. **45.0000 degrees when standardised** (Calculation 4.1): **always, in two variables**.
9. Fraction **0.7899 matches $(1 + \rho)/2$** (Observation 4.3): **the formula closes**.
10. **Standardising is a choice too** (Observation 4.5): **it is not neutral**.

Closing: Principal component analysis is described as extracting the main direction in the data——but that direction is not fixed by the data alone. PCA looks for the direction of greatest variance, and “greatest” is measured by the size of the numbers——yet the size of numbers changes with units: 170 cm is also 1.70 m and 1700 mm, one length producing variances that differ by 10^6 . So on one sample the direction moves 85 degrees——holding weight in kg and changing only the height unit, the first component points at **55.6503 degrees in cm**, **89.7345 in m**, and **4.1934 in mm**. The m and mm answers differ by 85.5411 degrees — nearly orthogonal. 89.7345 is nearly the weight axis and 4.1934 nearly the height axis, so “the main variation is weight” and “it is height” swap on the choice of unit alone. The same sample, the same people, and the values were only multiplied. The explained fraction moves too——1.0000 in m and 0.9898 in mm. Both read as “the first component explains nearly everything”, and the axis being explained is different. A high fraction does not guarantee that the direction means anything: push the units far enough and it approaches 1 as closely as you like. The eigenvalues (98.318922 against 6312.309583) are not comparable either, since their units differ. Standardising settles it——with both variables standardised the direction is exactly **45.0000 degrees** and the fraction **0.7899**. The two variances become equal, and the eigenvectors of a correlation matrix are always at 45 and 135 degrees in two variables. The fraction matches $(1 + \rho)/2$ for $\rho = 0.6$, so the formula closes. One thing separates them——whether the covariance or the correlation matrix is used; not the

data. Where units are meaningful, the covariance matrix; where they are mixed, the correlation matrix — not “which is right” but “what should weigh more”. By the criterion of Paper 300 they have distinct roots: the structure comes from the sample, the scale from the analyst, and the result is their product. To be said honestly——standardising is itself a choice — the decision to give every variable equal weight, which is not neutral. Where a high-variance variable ought to weigh more, the covariance matrix is right. The choice does not vanish; it changes place. One last thing——when one says “the data speak”, someone chose the scale in which they speak. Changing a unit alone turns the first component by 85 degrees, so that direction is not a property of the data alone. Report an analysis without its scale and it cannot be reproduced.

References

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. Yamagishi, Y. (2026). *The Three Hundredth Census: Whether Two Things Share a Root Is Decidable* [Paper 300]. Zenodo. 10.5281/zenodo.22171914
7. Yamagishi, Y. (2026). *The Current Flows and Does No Work* [Paper 356]. Zenodo. 10.5281/zenodo.22697124
8. Yamagishi, Y. (2026). *The “Depth” of Depth of Field Is a Convention, Not a Measurement* [Paper 358]. Zenodo. 10.5281/zenodo.22697216
9. Yamagishi, Y. (2026). *Halving the Area Does Not Halve the Species — And Halving the Species Takes Cutting the Area to 6.25 Per Cent* [Paper 372]. Zenodo. 10.5281/zenodo.22698241

— *From thinking to syncing.* —

Yuuki YAMAGISHI (山岸 勇輝)

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