Spectral Decomposition, the Simple Visual Story

An ultra-simple companion to Lesson 3 (PCA and EFA)

Prof. Dr. Jörg Osterrieder

School of Engineering and Computer Sciences

A Puzzle Before We Start

You have a cloud of two-dimensional data. The two variables are strongly correlated, so the cloud is a stretched, tilted oval.

Which single direction through that cloud carries the most information, and how much is left over for the rest?

Spectral decomposition answers exactly this, and it is the engine inside PCA. We will see it with one tiny 2 by 2 matrix and no proofs.

Before any maths: in a tilted oval cloud, where would you draw the most informative line?

1. Start With a Tiny Correlation Matrix

A correlation matrix C records how each variable moves with the others (1 on the diagonal, the correlation off-diagonal):

$$C = \begin{pmatrix} 1.0 & 0.8 \\ 0.8 & 1.0 \end{pmatrix}$$

- The diagonal 1.0: each variable correlates perfectly with itself.
- The off-diagonal 0.8: the two variables move together strongly.

Spectral decomposition rewrites C as a set of *directions* (eigenvectors) plus a *strength* for each (eigenvalues).

What does the off-diagonal 0.8 **tell you about the shape of the data cloud?**

2. What We Are Looking For

We want two things from C:

- **Eigenvectors**: special directions where the data only stretches, it does not bend.
- **Eigenvalues** (λ): how much the data spreads along each of those directions.

Predict: the two variables are positively correlated (0.8). Roughly which way will the longest, most-spread direction point?

2. What We Are Looking For

We want two things from C:

- **Eigenvectors**: special directions where the data only stretches, it does not bend.
- **Eigenvalues** (λ): how much the data spreads along each of those directions.

Predict: the two variables are positively correlated (0.8). Roughly which way will the longest, most-spread direction point? Along the 45° line (the $(1, 1)$ direction), because when both variables rise together, the cloud stretches diagonally.

Why should the most-spread direction follow the way the two variables move together?

3. Solve It With One Line of Arithmetic

The eigenvalues are the λ that make $\det(C - \lambda I) = 0$:

$$\det\begin{pmatrix} 1-\lambda & 0.8 \\ 0.8 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - 0.64 = 0$$

So $(1-\lambda)^2 = 0.64$, giving $1 - \lambda = \pm 0.8$. *Predict which is bigger before the reveal.*

3. Solve It With One Line of Arithmetic

The eigenvalues are the λ that make $\det(C - \lambda I) = 0$:

$$\det\begin{pmatrix} 1-\lambda & 0.8 \\ 0.8 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - 0.64 = 0$$

So $(1-\lambda)^2 = 0.64$, giving $1 - \lambda = \pm 0.8$. *Predict which is bigger before the reveal.*

$$\lambda_1 = 1.8 \qquad \text{and} \qquad \lambda_2 = 0.2$$

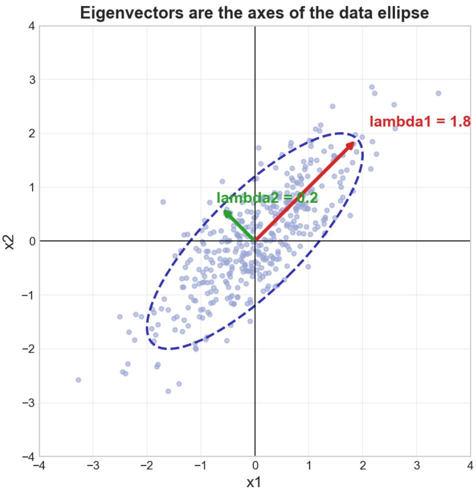
The eigenvectors come out as $(0.71, 0.71)$ for λ_1 and $(0.71, -0.71)$ for λ_2 (the 45° and anti-diagonal lines).

From $(1-\lambda)^2 = 0.64$**, how do you get the two eigenvalues** 1.8 **and** 0.2**?**

The eigenvectors are the axes of the data ellipse. The eigenvalue is how far the cloud stretches along each axis:

- Long axis: $\lambda_1 = 1.8$, direction $(0.71, 0.71)$.
- Short axis: $\lambda_2 = 0.2$, direction $(0.71, -0.71)$.

The big eigenvalue points where the data is most informative; the small one is the leftover, nearly flat direction.



Which arrow is longer, and what does its length actually represent?

5. Why This Is Exactly PCA

Principal Component Analysis is just the spectral decomposition of the correlation (or covariance) matrix:

- **Eigenvectors** are the principal component directions.
- **Eigenvalues** are the variance captured by each component.

Predict: what fraction of the total information sits on the first component here?

5. Why This Is Exactly PCA

Principal Component Analysis is just the spectral decomposition of the correlation (or covariance) matrix:

- **Eigenvectors** are the principal component directions.
- **Eigenvalues** are the variance captured by each component.

Predict: what fraction of the total information sits on the first component here? Total spread $= 1.8 + 0.2 = 2.0$. So PC1 holds $1.8/2.0 = \mathbf{90}\%$ and PC2 holds $0.2/2.0 = \mathbf{10}\%$. Keeping PC1 alone loses only 10%.

Spectral decomposition of a 2x2 correlation matrix

C =
```
[ 1.0   0.8 ]
[ 0.8   1.0 ]
```

```
det(C - lambda I) = (1 - lambda)^2 - 0.64 = 0
```

lambda_1 = 1.8 (90% of variance) eigenvector ~ (0.71, 0.71)

lambda_2 = 0.2 (10% of variance) eigenvector ~ (0.71, -0.71)

Eigenvectors give the directions, eigenvalues give the spread.

If PC1 carries 90% **of the variance, why is dropping PC2 a safe simplification here?**

The One Big Idea

Spectral decomposition splits a matrix into **directions** (eigenvectors) and **strengths** (eigenvalues). PCA is that split applied to your data's correlation matrix.

- $C = \begin{pmatrix} 1.0 & 0.8 \\ 0.8 & 1.0 \end{pmatrix}$ gives $\lambda_1 = 1.8$ and $\lambda_2 = 0.2$ from $(1 - \lambda)^2 = 0.64$.
- Eigenvectors $(0.71, 0.71)$ and $(0.71, -0.71)$ are the ellipse axes.
- λ_1 is PC1 (90% of variance), λ_2 is PC2 (10%).

Could you now explain, in one sentence, why eigenvalues equal the variance per component?