

Rotating the Principal Components¹

STA431 Spring 2023

¹See last slide for copyright information.

Summary

- Rotation is what makes exploratory factor analysis results understandable.
- R's stand-alone `varimax` function can also be used to rotate principal components.
- The result is a set of uncorrelated linear combinations of the variables that explain exactly the same amount of variance as the original components, but are easier to interpret.

Setting

- Standardized data vector \mathbf{z} is $k \times 1$
- $cov(\mathbf{z}) = \Sigma$
- $\Sigma = \mathbf{C}\mathbf{D}\mathbf{C}^\top$
- $\mathbf{y} = \mathbf{C}^\top \mathbf{z} \iff \mathbf{z} = \mathbf{C}\mathbf{y}$
- $cov(\mathbf{y}) = \mathbf{D}$

Standardize and Select Principal Components

Standardize first for convenience

- $cov(\mathbf{y}) = \mathbf{D}$
- Let $\mathbf{y}_2 = \mathbf{D}^{-\frac{1}{2}}\mathbf{y}$
- $cov(\mathbf{y}_2) = \mathbf{I}_k$

$$\begin{aligned}cor(\mathbf{d}, \mathbf{y}) &= cov(\mathbf{z}, \mathbf{y}_2) \\&= cov(\mathbf{z}, \mathbf{D}^{-\frac{1}{2}}\mathbf{y}) \\&= cov(\mathbf{z}, \mathbf{D}^{-\frac{1}{2}}\mathbf{C}^\top \mathbf{z}) \\&= cov(\mathbf{z}, \mathbf{z}) \left(\mathbf{D}^{-\frac{1}{2}}\mathbf{C}^\top \right)^\top \\&= \mathbf{\Sigma C D}^{-\frac{1}{2}} \\&= \mathbf{C D C}^\top \mathbf{C D}^{-\frac{1}{2}} \\&= \mathbf{C D}^{\frac{1}{2}}\end{aligned}$$

$cor(\mathbf{z}, \mathbf{y}) = \mathbf{CD}^{\frac{1}{2}}$ is a good matrix

- Square all the elements and get components of variance.
- Squared correlations add to one for each row.
- Squared correlations add to eigenvalues for each column.

Squared correlations add to one for each row.

$$\text{cor}(\mathbf{z}, \mathbf{y}) = \mathbf{CD}^{\frac{1}{2}}$$

Look at diagonal elements of

$$\begin{aligned}\mathbf{CD}^{\frac{1}{2}} \left(\mathbf{CD}^{\frac{1}{2}} \right)^{\top} &= \mathbf{CD}^{\frac{1}{2}} \mathbf{D}^{\frac{1}{2}} \mathbf{C}^{\top} \\ &= \mathbf{CDC}^{\top} \\ &= \mathbf{\Sigma} = \text{cov}(\mathbf{z})\end{aligned}$$

Diagonal elements are all ones.

Squared correlations add to eigenvalues for each column

$$\text{cor}(\mathbf{z}, \mathbf{y}) = \mathbf{CD}^{\frac{1}{2}}$$

Look at diagonal elements of

$$\begin{aligned}\left(\mathbf{CD}^{\frac{1}{2}}\right)^{\top} \mathbf{CD}^{\frac{1}{2}} &= \mathbf{D}^{\frac{1}{2}} \mathbf{C}^{\top} \mathbf{CD}^{\frac{1}{2}} \\ &= \mathbf{D}^{\frac{1}{2}} \mathbf{D}^{\frac{1}{2}} \\ &= \mathbf{D}\end{aligned}$$

Eigenvalues.

Select First p principal Components

Probably those with eigenvalues greater than one

$$\begin{aligned}\mathbf{z} &= \mathbf{C}\mathbf{y} \\ &= \mathbf{C}\mathbf{D}^{\frac{1}{2}}\mathbf{D}^{-\frac{1}{2}}\mathbf{y} \\ &= \underbrace{\mathbf{C}\mathbf{D}^{\frac{1}{2}}}_{k \times k} \underbrace{\mathbf{y}_2}_{k \times 1} \\ &= \left(\underbrace{\mathbf{L}}_{k \times p} \mid \underbrace{\mathbf{M}}_{k \times (k-p)} \right) \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix} \begin{matrix} \leftarrow p \times 1 \\ \leftarrow (k-p) \times 1 \end{matrix} \\ &= \mathbf{L}\mathbf{f} + \mathbf{M}\mathbf{g} \\ &= \mathbf{L}\mathbf{f} + \mathbf{e}\end{aligned}$$

$$\mathbf{z} = \mathbf{L}\mathbf{f} + \mathbf{e}$$

- It looks like a factor analysis model.
- \mathbf{f} contains the first p principal components, standardized.
- $\text{cov}(\mathbf{f}) = \mathbf{I}_p$.
- \mathbf{L} contains the first p columns of $\text{cor}(\mathbf{d}, \mathbf{y}) = \mathbf{CD}^{\frac{1}{2}}$.
- Recalling

$$\begin{aligned}\mathbf{z} &= (\mathbf{L} \mid \mathbf{M}) \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix} \\ &= \mathbf{L}\mathbf{f} + \mathbf{M}\mathbf{g} \\ &= \mathbf{L}\mathbf{f} + \mathbf{e},\end{aligned}$$

have $\text{cov}(\mathbf{f}, \mathbf{e}) = \mathbf{O}$.

- Results for factor analysis apply:
 - Components of variance explained by \mathbf{f} are squared correlations.
 - Communalities (explained variance of each variable) are not affected by rotation.

$$\begin{aligned}\mathbf{z} &= \mathbf{L}\mathbf{f} + \mathbf{e} \\ &= \mathbf{L}\mathbf{R}^\top\mathbf{R}\mathbf{f} + \mathbf{e} \\ &= (\mathbf{L}\mathbf{R}^\top)(\mathbf{R}\mathbf{f}) + \mathbf{e} \\ &= \mathbf{L}_2\mathbf{f}' + \mathbf{e},\end{aligned}$$

where \mathbf{L}_2 is the “rotated factor matrix,” and \mathbf{f}' are the rotated principal components.

$$\mathbf{z} = \mathbf{L}_2 \mathbf{f}' + \mathbf{e}$$

- $cov(\mathbf{f}') = \mathbf{I}_p$, so the rotated components are still uncorrelated.
- $cov(\mathbf{z}, \mathbf{f}') = \mathbf{L}_2$ is a matrix of correlations.
- You can examine $\hat{\mathbf{L}}_2$ to determine what the rotated factors *mean* in terms of the original variables.
- Rotation affects how much variance each component explains, but not the total amount of variance explained.
- Rotation does *not* affect the amount of explained variance for each variable.

In Practice it's Very Simple

- Extract sample principal components. and decide how many to keep.
- Put the ones you decide to keep in $\mathbf{Y}_{n \times p}$.
- Apply a varimax rotation to estimated $cor(\mathbf{D}, \mathbf{Y})$. This is \hat{L}_2 .
- If you like the result,
 - Standardize the principal components in \mathbf{Y} , using n in the denominator. Call the result \mathbf{W} . The rows of \mathbf{W} are approximately $\mathbf{f}_1, \dots, \mathbf{f}_n$.
 - Compute $\mathbf{W}\mathbf{R}^\top$, where \mathbf{R}^\top is the rotation matrix located by `varimax`.
 - The rows are the rotated sample principal components.

Copyright Information

This slide show was prepared by **Jerry Brunner**, Department of Statistical Sciences, University of Toronto. It is licensed under a **Creative Commons Attribution - ShareAlike 3.0 Unported License**. Use any part of it as you like and share the result freely. The L^AT_EX source code is available from the course website:

<http://www.utstat.toronto.edu/brunner/oldclass/431s23>