Eigenvalues, Eigenvectors, and Their Uses

James H. Steiger

Department of Psychology and Human Development
Vanderbilt University

Eigenvalues, Eigenvectors, and Their Uses

Introduction

- In this module, we explore the properties of eigenvalues, eigenvectors, and some of their key uses in statistics.
- We begin by defining eigenvalues and eigenvectors, and then we demonstrate some of their key mathematical properties.
- In conclusion, we show some key uses:
  1. Matrix factorization
  2. Least squares approximation
  3. Calculation of symmetric powers of symmetric matrices

Defining Eigenvalues and Eigenvectors

Definition (Eigenvalues and Eigenvectors)

For a square matrix $\mathbf{A}$, a scalar c and a vector $\mathbf{v}$ are an eigenvalue and associated eigenvector, respectively, if and only if they satisfy the equation,

$$\mathbf{A}\mathbf{v} = c\mathbf{v} \tag{1}$$

There are infinitely many solutions to Equation 1 unless some identification constraint is placed on the size of vector $\mathbf{v}$. For example for any c and $\mathbf{v}$ satisfying the equation, $c/2$ and $2\mathbf{v}$ must also satisfy the same equation. Consequently in eigenvectors are assumed to be "normalized," i.e., satisfy the constraint that $\mathbf{v}'\mathbf{v} = 1$. Eigenvalues c_i are roots to the determinantal equation

$$|\mathbf{A} - c\mathbf{I}| = 0 \tag{2}$$

Key Properties of Eigenvalues and Eigenvectors I

Here are some key properties of eigenvalues and eigenvectors. For $n \times n$ matrix $\mathbf{A}$ with eigenvalues c_i and associated eigenvectors $\mathbf{v}_i$,

1. $$\mathrm{Tr}\left(\mathbf{A}\right) = \sum_{i=1}^{n} c_i$$

2. $$|\mathbf{A}| = \prod_{i=1}^{n} c_i$$

3. Eigenvalues of a symmetric matrix with real elements are all real.

4. Eigenvalues of a positive definite matrix are all positive.

Key Properties of Eigenvalues and Eigenvectors II

5. If a $n \times n$ symmetric matrix $\mathbf{A}$ is positive semidefinite and of rank r, it has exactly r positive eigenvalues and $p - r$ zero eigenvalues.
6. The nonzero eigenvalues of the product $\mathbf{AB}$ are equal to the nonzero eigenvalues of $\mathbf{BA}$. Hence the traces of $\mathbf{AB}$ and $\mathbf{BA}$ are equal.
7. The characteristic roots of a diagonal matrix are its diagonal elements.
8. The scalar multiple $b\mathbf{A}$ has eigenvalue bc_i with eigenvector $\mathbf{v}_i$.

   Proof: $\mathbf{A}\mathbf{v}_i = c_i\mathbf{v}_i$ implies immediately that $(b\mathbf{A})\mathbf{v}_i = (bc_i)\mathbf{v}_i$.
9. Adding a constant b to every diagonal element of $\mathbf{A}$ creates a matrix $\mathbf{A} + b\mathbf{I}$ with eigenvalues $c_i + b$ and associated eigenvectors $\mathbf{v}_i$.

   Proof: $(\mathbf{A} + b\mathbf{I})\mathbf{v}_i = \mathbf{A}\mathbf{v}_i + b\mathbf{v}_i = c_i\mathbf{v}_i + b\mathbf{v}_i = (c_i + b)\mathbf{v}_i$.

Key Properties of Eigenvalues and Eigenvectors III

10. $\mathbf{A}^m$ has c_i^m as an eigenvalue, and $\mathbf{v}_i$ as its eigenvector.

    Proof. Consider $\mathbf{A}^2\mathbf{v}_i = \mathbf{A}(\mathbf{A}\mathbf{v}_i) = \mathbf{A}(\mathbf{c}_i\mathbf{v}_i) = c_i(\mathbf{A}\mathbf{v}_i) = c_i c_i \mathbf{v}_i = c_i^2\mathbf{v}_i$. The general case follows by induction.

11. $\mathbf{A}^{-1}$, if it exists, has $1/c_i$ as an eigenvalue, and $\mathbf{v}_i$ as its eigenvector.

    Proof. $\mathbf{A}\mathbf{v}_i = c_i\mathbf{v}_i = \mathbf{v}_i c$. $\mathbf{A}^{-1}\mathbf{A}\mathbf{v}_i = \mathbf{v}_i = \mathbf{A}^{-1}\mathbf{v}_i c_i$. $\mathbf{v}_i = \mathbf{A}^{-1}\mathbf{v}_i c_i = c_i\mathbf{A}^{-1}\mathbf{v}_i$. So $(1/c_i)\mathbf{v}_i = \mathbf{A}^{-1}\mathbf{v}_i$.

Key Properties of Eigenvalues and Eigenvectors IV

12. For symmetric $\mathbf{A}$, for distinct eigenvalues c_i, c_j with associated eigenvectors $\mathbf{v}_i$, $\mathbf{v}_j$, we have $\mathbf{v}_i'\mathbf{v}_j = 0$.

    Proof: $\mathbf{A}\mathbf{v}_i = c_i\mathbf{v}_i$, and $\mathbf{A}\mathbf{v}_j = c_j\mathbf{v}_j$. So $\mathbf{v}_j'\mathbf{A}\mathbf{v}_i = c_i\mathbf{v}_j'\mathbf{v}_i$ and $\mathbf{v}_i'\mathbf{A}\mathbf{v}_j = c_j\mathbf{v}_i'\mathbf{v}_j$. But, since a bilinear form is a scalar, it is equal to its transpose, and, remembering that $\mathbf{A} = \mathbf{A}'$, $\mathbf{v}_j'\mathbf{A}\mathbf{v}_i = \mathbf{v}_i'\mathbf{A}\mathbf{v}_j$. So $c_i\mathbf{v}_j'\mathbf{v}_i = c_j\mathbf{v}_i'\mathbf{v}_j = c_j\mathbf{v}_j'\mathbf{v}_i$. If c_i and c_j are different, this implies $\mathbf{v}_j'\mathbf{v}_i = 0$.

13. Eckart-Young Decomposition. For any real, symmetric $\mathbf{A}$, there exists a $\mathbf{V}$ such that $\mathbf{V}'\mathbf{A}\mathbf{V} = \mathbf{D}$,where $\mathbf{D}$ is diagonal. Moreover, any real symmetric matrix $\mathbf{A}$ can be written as $\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}'$,where $\mathbf{V}$ contains the eigenvectors $\mathbf{v}_i$ of $\mathbf{A}$ in order in its columns, and D contains the eigenvalues c_i of $\mathbf{A}$ in the ith diagonal position. (Proof?)

Key Properties of Eigenvalues and Eigenvectors V

14. Best Rank-m Least Squares Approximation. Recall from our reading that a set of vectors is linearly independent if no vector in the set is a linear combination of the others, and that the rank of a matrix is the the (largest) number of rows and columns that exhibit linear independence. In general, if we are approximating one symmetric matrix with another, matrices of higher rank (being less restricted) can do a better job of approximating a full-rank matrix $\mathbf{A}$ than matrices of lower rank. Suppose that the eigenvectors and eigenvalues of symmetric matrix $\mathbf{A}$ are ordered in the matrices $\mathbf{V}$ and $\mathbf{D}$ in descending order, so that the first element of $\mathbf{D}$ is the largest eigenvalue of $\mathbf{A}$, and the first column of $\mathbf{V}$ is its corresponding eigenvector. Define $\mathbf{V}^*$ as the first m columns of $\mathbf{V}$, and $\mathbf{D}^*$ as an $m \times m$ diagonal matrix with the corresponding m eigenvalues as diagonal entries. Then $\mathbf{V}^*\mathbf{D}^*\mathbf{V}^{*\prime}$ is a matrix of rank m that is the best possible (in the least squares sense) rank m approximation of $\mathbf{A}$.

Key Properties of Eigenvalues and Eigenvectors VI

15. Consider all possible "normalized quadratic forms in $\mathbf{A}$," i.e.,

$$q(\mathbf{x}_i) = \mathbf{x}_i' \mathbf{A} \mathbf{x}_i \tag{3}$$

with $\mathbf{x}_i' \mathbf{x}_i = 1$.

The maximum of all quadratic forms is achieved with $\mathbf{x}_i = \mathbf{v}_1$, where $\mathbf{v}_1$ is the eigenvector corresponding to the largest eigenvalue of $\mathbf{A}$. The minimum is achieved with $\mathbf{x}_i = \mathbf{v}_m$, the eigenvector corresponding to the smallest eigenvalue of $\mathbf{A}$. The maxima and minima are the largest and smallest eigenvalues, respectively.

Applications of Eigenvalues and Eigenvectors

Powers of a Diagonal Matrix

Eigenvalues and eigenvectors have widespread practical application in multivariate statistics. In this section, we demonstrate a few such applications. First, we deal with the notion of *matrix factorization*.

Definition (Powers of a Diagonal Matrix)

Diagonal matrices act much more like scalars than most matrices do. For example, we can define fractional powers of diagonal matrices, as well as positive powers. Specifically, if diagonal matrix $\mathbf{D}$ has diagonal elements d_i, the matrix $\mathbf{D}^x$ has elements d_i^x. If x is negative, it is assumed $\mathbf{D}$ is positive definite. With this definition, the powers of $\mathbf{D}$ behave essentially like scalars. For example, $\mathbf{D}^{1/2}\mathbf{D}^{1/2} = \mathbf{D}$.

Applications of Eigenvalues and Eigenvectors

Powers of a Diagonal Matrix

Example (Powers of a Diagonal Matrix)

Suppose we have

$$\mathbf{D} = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$$

Then

$$\mathbf{D}^{1/2} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

Applications of Eigenvalues and Eigenvectors

Matrix Factorization

Example (Matrix Factorization)

Suppose you have a variance-covariance matrix $\boldsymbol{\Sigma}$ for some statistical population. Assuming $\boldsymbol{\Sigma}$ is positive semidefinite, then (from Result 5), it can be written in the form $\boldsymbol{\Sigma} = \mathbf{VDV}' = \mathbf{FF}'$, where $\mathbf{F} = \mathbf{VD}^{1/2}$. $\mathbf{F}$ is called a "Gram-factor of $\boldsymbol{\Sigma}$."

- Gram-factors are not, in general, uniquely defined.
- For example, suppose $\boldsymbol{\Sigma} = \mathbf{FF}'$. Then consider any orthogonal matrix $\mathbf{T}$, such that $\mathbf{TT}' = \mathbf{T}'\mathbf{T} = \mathbf{I}$.
- There are infinitely many orthogonal matrices $\mathbf{T}$ of order 2×2 and higher. For any such matrix $\mathbf{T}$, we have $\boldsymbol{\Sigma} = \mathbf{FTT}'\mathbf{F}' = \mathbf{F}^*\mathbf{F}^{*\prime}$, where $\mathbf{F}^* = \mathbf{FT}$.

Applications of Eigenvalues and Eigenvectors

Matrix Factorization

- Gram-factors have some significant applications.
- For example, in the field of random number generation, it is relatively easy to generate pseudo-random numbers that mimic p variables that are independent with zero mean and unit variance.
- But suppose we wish to mimic p variables that are not independent, but have variance-covariance matrix $\boldsymbol{\Sigma}$?
- The next example describes one method for doing this.

Example (Simulating Nonindependent Random Numbers)

Given $p \times 1$ random vector $\mathbf{x}$ having variance-covariance matrix $\mathbf{I}$. Let $\mathbf{F}$ be a Gram-factor of $\boldsymbol{\Sigma} = \mathbf{F}\mathbf{F}'$. Then $\mathbf{y} = \mathbf{F}\mathbf{x}$ will have variance-covariance matrix $\boldsymbol{\Sigma}$.

Symmetric Powers of a Symmetric Matrix

In certain intermediate and advanced derivations in matrix algebra, reference is made to "symmetric powers" of a symmetric matrix $\mathbf{A}$, in particular the "symmetric square root" of $\mathbf{A}$, a symmetric matrix which, when multiplied by itself, yields $\mathbf{A}$.

Example (Symmetric Powers of a Symmetric Matrix)

When investigating properties of eigenvalues and eigenvectors, we pointed out that, for distinct eigenvalues of a symmetric matrix $\mathbf{A}$, the associated eigenvectors are orthogonal. Since the eigenvectors are normalized to have a sum of squares equal to 1, it follows that if we place the eigenvectors in a matrix $\mathbf{V}$, this matrix will be orthogonal, i.e. $\mathbf{VV}' = \mathbf{V}'\mathbf{V} = \mathbf{I}$. This fact allows us to create "symmetric powers" of a symmetric matrix very efficiently if we know the eigenvectors. (continued on next slide . . .)

Symmetric Powers of a Symmetric Matrix

Example (Symmetric Powers of a Symmetric Matrix)

For example, suppose you wish to create a symmetric matrix $\mathbf{A}^{1/2}$ such that $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$. Let diagonal matrix $\mathbf{D}$ contain the eigenvalues of $\mathbf{A}$ in proper order. Then $\mathbf{A} = \mathbf{VDV}'$, and it is easy to verify that $\mathbf{A}^{1/2} = \mathbf{VD}^{1/2}\mathbf{V}'$ has the required properties. To prove that $\mathbf{A}^{1/2}$ is symmetric, we need simply show that it is equal to its transpose, which is trivial (so long as you recall that any diagonal matrix is symmetric, and that the transpose of a product of several matrices is the product of the transposes in reverse order). That $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$ follows immediately by substitution, i.e.,

$$
\begin{aligned}
\mathbf{A}^{1/2}\mathbf{A}^{1/2} &= \mathbf{VD}^{1/2}\mathbf{V}'\mathbf{VD}^{1/2}\mathbf{V}' \\
&= \mathbf{VD}^{1/2}\left[\mathbf{V}'\mathbf{V}\right]\mathbf{D}^{1/2}\mathbf{V}' \\
&= \mathbf{VD}^{1/2}\left[\mathbf{I}\right]\mathbf{D}^{1/2}\mathbf{V}' \\
&= \mathbf{VD}^{1/2}\mathbf{D}^{1/2}\mathbf{V}' \\
&= \mathbf{VDV}'
\end{aligned}
$$

Some Eigenvalue-Eigenvector Calculations in R

Extracting Eigenvalues and Eigenvectors

- Extracting eigenvalues and eigenvectors in R is straightforward.
- Consider the following 3×3 correlation matrix.

```
> Rxx <- matrix(c(1, 0.5, 0.4, 0.5, 1, 0.3, 0.4, 0.3, 1), 3, 3)
> Rxx

     [,1] [,2] [,3]
[1,]  1.0  0.5  0.4
[2,]  0.5  1.0  0.3
[3,]  0.4  0.3  1.0
```

Some Eigenvalue-Eigenvector Calculations in R

Extracting Eigenvalues and Eigenvectors

- We can extract the eigenvalues and eigenvectors together in a structure with the `eigen` command:

```
> eigendecomp <- eigen(Rxx)
```

- We have saved the eigendecomposition results in the variable `eigendecomp`.
- It turns out the the eigenvalues are in this structure in a matrix called `vectors`, and the eigenvalues are in a vector called `values`.

Some Eigenvalue-Eigenvector Calculations in R

Extracting Eigenvalues and Eigenvectors

```
> V <- eigendecomp$vectors
> d <- eigendecomp$values
> V

          [,1]       [,2]       [,3]
[1,] 0.6215471 -0.1703679  0.7646267
[2,] 0.5813269 -0.5539601 -0.5959759
[3,] 0.5251078  0.8149252 -0.2452728

> d

[1] 1.8055810 0.7124457 0.4819732
```

Some Eigenvalue-Eigenvector Calculations in R

Constructing a Gram Factor

- Recall that a Gram factor of $\mathbf{R_{xx}}$, $\mathbf{F}$, can be constructed so that $\mathbf{R_{xx}} = \mathbf{FF}'$. Since $\mathbf{R_{xx}} = \mathbf{VDV}'$, an obvious choice for $\mathbf{F}$ is $\mathbf{F} = \mathbf{VD}^{1/2}$. Below we compute $\mathbf{F}$ and verify that it is indeed a Gram-factor of $\mathbf{R_{xx}}$.

```
> F = V %*% diag(sqrt(d))
> F

          [,1]       [,2]       [,3]
[1,] 0.8351847 -0.1438016  0.5308366
[2,] 0.7811400 -0.4675783 -0.4137521
[3,] 0.7055974  0.6878498 -0.1702789

> F %*% t(F)

     [,1] [,2] [,3]
[1,]  1.0  0.5  0.4
[2,]  0.5  1.0  0.3
[3,]  0.4  0.3  1.0
```

Some Eigenvalue-Eigenvector Calculations in R

Constructing a Cholesky Factor

- A Cholesky factor of $\mathbf{R_{xx}}$ is a Gram-factor that is also a triangular matrix.
- Note: The `chol` function in R departs from the standard convention in that it returns the upper triangular factor $\mathbf{F}'$ rather than the lower triangular factor $\mathbf{F}$.
- So, to create a Gram factor $\mathbf{F}$ such that $\mathbf{R_{xx}} = \mathbf{FF}'$, one must transpose the matrix returned by the `chol` function.

Some Eigenvalue-Eigenvector Calculations in R

Constructing a Cholesky Factor

- Here we demonstrate how to extract a Cholesky factor of $\mathbf{R_{xx}}$.
- We then demonstrate that the matrix we extract is in fact a Gram factor of $\mathbf{R_{xx}}$.

```
> F.c = t(chol(Rxx))
> F.c

     [,1]      [,2]      [,3]
[1,]  1.0 0.0000000 0.0000000
[2,]  0.5 0.8660254 0.0000000
[3,]  0.4 0.1154701 0.9092121

> F.c %*% t(F.c)

     [,1] [,2] [,3]
[1,]  1.0  0.5  0.4
[2,]  0.5  1.0  0.3
[3,]  0.4  0.3  1.0
```

Some Eigenvalue-Eigenvector Calculations in R

Constructing a Symmetric Power of a Symmetric Matrix

- Here is a function for computing symmetric powers of a symmetric matrix.

```
> sympower <- function(x, pow) {
+     edecomp <- eigen(x)
+     roots <- edecomp$val
+     v <- edecomp$vec
+     d <- roots^pow
+     if (length(roots) == 1)
+         d <- matrix(d, 1, 1) else d <- diag(d)
+     sympow <- v %*% d %*% t(v)
+     sympow
+ }
```

Some Eigenvalue-Eigenvector Calculations in R

Constructing a Symmetric Power of a Symmetric Matrix

- Below we compute $\mathbf{R}_{\mathbf{xx}}^{1/2}$, the symmetric square root of $\mathbf{R}_{\mathbf{xx}}$.

```
> Rxx.half <- sympower(Rxx, 1/2)
> Rxx.half

          [,1]      [,2]      [,3]
[1,] 0.9494977 0.2488098 0.1911747
[2,] 0.2488098 0.9597037 0.1306236
[3,] 0.1911747 0.1306236 0.9728256

> Rxx.half %*% Rxx.half

     [,1] [,2] [,3]
[1,]  1.0  0.5  0.4
[2,]  0.5  1.0  0.3
[3,]  0.4  0.3  1.0
```