Data Visualization

Lecture 5
Mathamatics Visualization

Phuc Loi Luu, PhD
p.luu@garvan.org.au
luu.p.loi@googlemail.com

Content

Plot a mathmatics graph

- Introduction to Linear Algebra 7: Eigen Values and Eigen Vectors

Eigen Values and Eigen Vectors

Suppose we have a 2×2 matrix:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Then notice that if I have a vector

$$v = \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

then we have

$$A \cdot v = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Therefore, we have a relation:

$$A \cdot v = v.$$

In addition, suppose we have another vector:

$$u = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

then we have

$$A \cdot u = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}.$$

Therefore, we have another relation:

$$A \cdot u = 3 \cdot u.$$

For each matrix there are special vectors, such as u and v in this example, whose images under the matrix are scalar multiplications of themselves. In general we have the following form: For an $n \times n$ matrix A, there is a vector v and a scalar λ such that

$$A \cdot v = \lambda \cdot v.$$

These vectors are called **eigen vectors** of a matrix and these scalars are called **eigen values**.

Compute Eigen Values and Eigen Vectors

Definition 48 *Suppose we have an $n \times n$ matrix A. Then an* **eigen vector** *of A is a non-zero vector such that*

$$A \cdot v = \lambda \cdot v.$$

A scalar λ is called an **eigen value** *associate with the eigen vector v.*

In order to compute the eigen values and eigen vectors of a matrix, first we set up the matrix equation:

$$A \cdot v = \lambda \cdot v.$$

Then we bring the right hand side to the left-hand side:

$$A \cdot v - \lambda \cdot v = 0.$$

Then we simplify this equation:

$$(A - \lambda \cdot I_n) \cdot v = 0,$$

where I_n is the identity matrix of size n. This means that an non-zero vector V is in the null space of the matrix $(A - \lambda \cdot I_n)$. Since the matrix $(A - \lambda \cdot I_n)$ has non-zero vector in the null space, the dimension of the null space is greater than or equal to 1. Thus, the determinant of the matrix $(A - \lambda \cdot I_n)$ must be equal to 0. So we set

$$\det (A - \lambda \cdot I_n) = 0,$$

which is a polynomial in terms of one variable λ. This polynomial is called the **characteristic polynomial** of a matrix A. The roots of the characteristic polynomial of A are the eigen values. From each root λ_i of the characteristic polynomial, we find the null space of the matrix $(A - \lambda \cdot I_n)$. All elements in a basis for the null space of $(A - \lambda \cdot I_n)$ are eigen vectors associated with the eigen value λ_i.

Compute Eigen Values and Eigen Vectors

Example 141 *Suppose we have a 2×2 matrix:*

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Then, a characteristic polynomial of $(A - \lambda \cdot I_n)$ is

$$\det(A - \lambda \cdot I_n) = 0,$$

which is

$$\det\left(\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \lambda \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right)$$
$$= \det\left(\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}\right) = (2-\lambda)^2 - 1 = \lambda^2 - 4 \cdot \lambda + 3 = 0.$$

This polynomial has two roots: $\lambda_1 = 1$, $\lambda_2 = 3$. Now we want to find eigen vectors associated with $\lambda_1 = 1$, $\lambda_2 = 3$.

For $\lambda_1 = 1$, we substitute $\lambda = 1$ in the characteristic polynomial, then we have

$$C_1 = \begin{bmatrix} 2-1 & 1 \\ 1 & 2-1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Since a vector

$$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

is in the null space of the matrix C_1, a vector

$$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

spans the null space of the matrix C_1 and is an eigen vector associated with $\lambda_1 = 1$.

For $\lambda_1 = 3$, we substitute $\lambda = 3$ to the characteristic polynomial, then we have

$$C_2 = \begin{bmatrix} 2-3 & 1 \\ 1 & 2-3 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Since a vector

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is in the null space of the matrix C_2. Therefore a vector

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

spans the null space of the matrix C_2 and is an eigen vector associated with $\lambda_2 = 3$.

In R: compute Eigen Values and Eigen Vectors

In R, we can use the eigen() function to compute eigen values and eigen vectors of a matrix. For Example 141, consider a matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

First we define the matrix in R:

```
A <- matrix(c(2, 1, 1, 2), 2, 2)
```

Then we call the eigen() function:

```
eigen(A)
```

R returns the eigen values and eigen vectors associated with them:

```
> eigen(A)
eigen() decomposition
$values
[1] 3 1

$vectors
          [,1]       [,2]
[1,] 0.7071068 -0.7071068
[2,] 0.7071068  0.7071068
```

Here, the first column of \$vectors is an eigen vector associated with the eigen value equal to 3, which is the first value of \$values. The second column of \$vectors is an eigen vector associated with the eigen value equal to 1, which is the second value of \$values. Note that these vectors are already "normalized", i.e., these length is equal to 1.

There are some theorems which help us to compute its eigen values and eigen vectors. We will discuss some nice properties of eigen values and eigen vectors of a matrix.

Properties of Eigen Values and Eigen Vectors

Theorem 6.1 *Eigen values of a triangular matrix are the entries on the diagonal of the matrix.*

Example 142 *Suppose we have a triangular matrix:*

$$A = \begin{bmatrix} 2 & 1 & 3 & -1 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Then the eigen values of the matrix A are 2, 1, -1.

Example 143 *Suppose we have a triangular matrix:*

$$A = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Then the eigen values of the matrix A are $3, -2$, 0, -1.

Theorem 6.2 *Suppose we have an $n \times n$ matrix A. A is invertible if and only if 0 is not an eigen value of A.*

Example 144 *From Example 141, suppose we have a 2×2 matrix:*

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

We computed its eigen values from the characteristic polynomial and they are $1, 3$, which are not 0, thus A is invertible.

Example 145 *From Example 142, suppose we have a triangular matrix:*

$$A = \begin{bmatrix} 2 & 1 & 3 & -1 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Then the eigen values of the matrix A are 2, 1, -1 *which are not equal to* 0. *Thus A is invertible.*

Properties of Eigen Values and Eigen Vectors

Example 146 *From Example 143, suppose we have a triangular matrix:*

$$A = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Then, the eigen values of the matrix A are $3, -2, 0, -1$. Since 0 is an eigen value, A is not invertible.

Theorem 6.3 *Suppose $v_1, v_2, \ldots, v_r$ are eigen vectors of a matrix A associated with the distinct eigen values $\lambda_1, \lambda_2, \ldots, \lambda_r$ of A. Then $v_1, v_2, \ldots, v_r$ are linearly independent.*

Example 147 *From Example 141, suppose we have a 2×2 matrix:*

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Its eigen values are $1, 3$, and their corresponding eigen vectors are

$$v_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \; v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

which are linearly independent.

Theorem 6.4 *Suppose $v_1, v_2, \ldots, v_r$ are eigen vectors of a symmetric matrix A associated with the distinct eigen values $\lambda_1, \lambda_2, \ldots, \lambda_r$ of A. Then $v_1, v_2, \ldots, v_r$ are linearly independent and they are orthogonal to each other.*

Example 148 *From Example 141, suppose we have a 2×2 matrix:*

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Its eigen values are $1, 3$, and their corresponding eigen vectors are

$$v_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \; v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

which are linearly independent. Also notice that

$$\langle v_1, v_2 \rangle = \langle \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rangle = (-1) \cdot 1 + 1 \cdot 1 = (-1) + 1 = 0.$$

Exercises

Exercise 6.1 *Compute the eigen values and eigen vectors of the following matrices without any help from a computer. Also check if they are invertible.*

1. $$\begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix}.$$

2. $$\begin{bmatrix} -1 & 3 \\ 1 & 5 \end{bmatrix}.$$

3. $$\begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}.$$

4. $$\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}.$$

5. $$\begin{bmatrix} 1 & 3 & 0 \\ 1 & 1 & 1 \\ -1 & 1 & 0 \end{bmatrix}.$$

6. $$\begin{bmatrix} 1 & 3 & 0 \\ 1 & 1 & 1 \\ 2 & 3 & 0 \end{bmatrix}.$$

Lab Exercise 175 *Compute the eigen values and eigen vectors of the following matrices with the eigen() function.*

1. $$\begin{bmatrix} -3 & 0 & 0 & 0 \\ 3 & -3 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 3 & -3 & -3 & -2 \end{bmatrix}$$

2. $$\begin{bmatrix} -1 & 0 & 1 & 3 \\ 0 & 0 & 2 & 6 \\ 1 & 2 & -1 & -2 \\ 3 & 6 & -2 & -1 \end{bmatrix}$$

3. $$\begin{bmatrix} -3 & 3 & 3 & 0 & -3 \\ 3 & 0 & 2 & -2 & 1 \\ 3 & 2 & 3 & 2 & -2 \\ 0 & -2 & 2 & -3 & 3 \\ -3 & 1 & -2 & 3 & 3 \end{bmatrix}$$

4. $$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & -3 & 2 & 0 & 0 \\ 0 & -3 & -2 & 0 & -3 \\ 0 & 0 & -2 & 0 & 0 \\ 2 & -2 & 1 & 0 & 0 \end{bmatrix}$$

Application in PCA

For Practical Applications in this section, we will discuss how we can compute the first and second principal components of the data set. Going back to the working example for this section, we generated 1000 data points under the bivariate normal distribution with

$$\mu = \begin{bmatrix} 5 \\ 5 \end{bmatrix}, \sigma = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix},$$

where μ is the mean and σ is the **covariance matrix** which defines the standard deviation for bivariate normal distribution. Note that this covariance matrix is symmetric. In order to define the standard deviation for a multivariate normal distribution, the covariance matrix has to be symmetric.

In real life, however, we do not observe μ and σ. We observe only the data set x which contains 1000 data points. Therefore we have to estimate them from the data points. Now we estimate a covariant matrix. Suppose x is the matrix whose rows represent each data point in the data set and whose columns represent each variable. Then we can estimate the covariance matrix $\hat{\sigma}$ by $x^T \cdot x$, where x^T is the transpose of the matrix x. $\hat{\sigma}$ is a $k \times k$ symmetric matrix where k is the number of random variables.

In R, we can estimate by:

```
y <- t(x) %*% x
```

Then we apply the eigen() function to compute the eigen values and eigen vectors of this estimated covariance matrix:

```
eigen(y)
```

This returns the following:

```
> eigen(y)
eigen() decomposition
$values
[1] 56444.411  1015.821

$vectors
           [,1]       [,2]
[1,] -0.7091532  0.7050544
[2,] -0.7050544 -0.7091532
```

Application in PCA

The eigen values of $\hat{\sigma}$ are 6444.411, 1015.821, and their eigen vectors are

$$\begin{bmatrix} -0.7091532 \\ -0.7050544 \end{bmatrix}, \begin{bmatrix} 0.7050544 \\ -0.7091532 \end{bmatrix},$$

respectively. Since the eigen values of $\hat{\sigma}$ are distinct and $\hat{\sigma}$ is symmetric, eigen vectors of $\hat{\sigma}$ are orthogonal and they are linearly independent. In fact, $\hat{\sigma}$ is always symmetric and if these points are random enough, then the eigen values are distinct, and all eigen vectors of $\hat{\sigma}$ are orthogonal to each other. This is one of the most important properties for PCA.

Now we look into the eigen vectors outputted from the eigen() function on $\hat{\sigma}$. The first principal component for this data set x is the eigen vector for the biggest eigen value of $\hat{\sigma}$. The second principal component is the eigen vector for the second biggest eigen value of $\hat{\sigma}$. In fact, the blue line in Figure 6.4 is defined by the eigen vector of the biggest eigen value of $\hat{\sigma}$, i.e., a set of all points defined by

$$\alpha \cdot \begin{bmatrix} -0.7091532 \\ -0.7050544 \end{bmatrix}$$

where α is any real number. In 6.4, the blue line is spanned by the first principal component and the pink line is spanned by the second principal component.

To check if these vectors are in fact the first and second principal components, we use the prcomp() function to verify:

```
> prcomp(d, center = FALSE)
Standard deviations (1, .., p=2):
[1] 7.516709 1.008384

Rotation (n x k) = (2 x 2):
          PC1        PC2
X1 -0.7091532  0.7050544
X2 -0.7050544 -0.7091532
```

This shows that the first and second principal components are

$$\begin{bmatrix} -0.7091532 \\ -0.7050544 \end{bmatrix}, \begin{bmatrix} 0.7050544 \\ -0.7091532 \end{bmatrix},$$

respectively.

Homework (https://github.com/friendly/matlib)

1. Properties of Eigenvalues and Eigenvectors
https://cran.r-project.org/web/packages/matlib/vignettes/eigen-ex1.html

2. Spectral Decomposition of Eigenvalues
https://cran.r-project.org/web/packages/matlib/vignettes/eigen-ex2.html