Topic 10: Eigenvalues and Eigenvectors

02-680: Essentials of Mathematics and Statistics

September 30, 2024

For a square matrix $A \in \mathbb{R}^{n \times n}$ for which $Ax = \lambda x$, we define $\lambda \in \mathbb{R}$ as an ***eigenvalue*** and $x \in \mathbb{R}^n \setminus \mathbf{0}$ as an ***eigenvector***. Together we call these ***eigenelements***.

The main idea is that the ***transformation*** of x by A results in a vector that points in the same direction of x, but scaled by λ.

Lets look at an example: let $A = \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}$, $x = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, and $\lambda = 5$.

$$\begin{array}{rcl} Ax & \stackrel{?}{=} & \lambda x \\ \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} & \stackrel{?}{=} & 5 \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ \begin{bmatrix} (3 \cdot 2) + (4 \cdot 1) \\ (2 \cdot 2) + (1 \cdot 1) \end{bmatrix} & \stackrel{?}{=} & \begin{bmatrix} (5 \cdot 2) \\ (5 \cdot 1) \end{bmatrix} \\ \begin{bmatrix} 10 \\ 5 \end{bmatrix} & = & \begin{bmatrix} 10 \\ 5 \end{bmatrix} \end{array}$$

but given only A, how can we find the eigenelements?

1 Finding Eigenelements

Lets break down the original equation:

$$\begin{array}{rcl} Ax & = & \lambda x \\ Ax & = & \lambda I_n x \\ Ax - \lambda I_n x & = & \lambda I_n x - \lambda I_n x \\ (A - \lambda I_n)x & = & 0 \end{array}$$

Since we know x cannot be $\mathbf{0}$ then $(A - \lambda I_n)$ but be 0 to satisfy the equation. This is actually a linear system on one variable! It turns out this only has non-trivial solutions when $(A - \lambda I_n)$ is singular (that is, it can't be invertible).

And remember we know that a matrix is singular when the determinant is 0:

$$|A - \lambda I_n| = 0.$$

We call this a ***characteristic equation***.

Lets look at the example from above, we can rewrite it this way:

$$\begin{aligned}
\begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} &= 0 \\
\begin{bmatrix} 3 - \lambda & 4 \\ 2 & 1 - \lambda \end{bmatrix} &= 0 \\
(3 - \lambda)(1 - \lambda) - (4)(2) &= 0 \\
(3 - 4\lambda + \lambda^2) - (8) &= 0 \\
\lambda^2 - 4\lambda - 5 &= 0 \\
(\lambda + 1)(\lambda - 5) &= 0
\end{aligned}$$

So we know that this matrix is singular when $\lambda = 5$ and -1. These are the eigenvalues of A, to find the eigenvectors we plug this back into the original equation and solve for x.

$$\begin{aligned}
\left(\begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix} - 5 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= 0 \\
\begin{bmatrix} -2 & 4 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= 0
\end{aligned}$$

We can use whatever method we want to solve the system but we will find that for

$$\lambda = 5, x = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Similarly for $\lambda = -1$ we get the system:

$$\begin{bmatrix} 4 & 4 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \rightarrow x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Note in both of these cases there is a trivial solution of $x = \mathbf{0}$ but this is disallowed by definition.

Aside

Consequence of above are

- That the determinant is the sum of the eigenvalues:

$$|A| = \sum_{i=1}^{n} \lambda_i$$

- The rank of A is equal to the number of non-zero eigenvalues
- If A is nonsingular, then $1/\lambda_i$ is an eigenvalue of A^{-1} with the same associated original eigenvector

A larger example. Find the eigenelements of $B = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 3 & -4 \\ 4 & 1 & -4 \end{bmatrix}$.

We first write down the characteristic equation:

$$\begin{aligned}
\begin{vmatrix} 1-\lambda & 1 & -1 \\ 2 & 3-\lambda & -4 \\ 4 & 1 & -4-\lambda \end{vmatrix} &= 0 \\
(1)\begin{vmatrix} 2 & -4 \\ 4 & -4-\lambda \end{vmatrix} - (3-\lambda)\begin{vmatrix} 1-\lambda & -1 \\ 4 & -4-\lambda \end{vmatrix} + (1)\begin{vmatrix} 1-\lambda & -1 \\ 2 & -4 \end{vmatrix} &= 0 \\
\text{(note we chose column 2 because it had more 1's)} & \\
\text{... details left to the reader ...} & \\
\lambda^3 - 7\lambda + 6 &= 0 \\
(\lambda-1)(\lambda-2)(\lambda+3) &= 0
\end{aligned}$$

Therefore our eigenvalues are $\lambda = 1, 2$, and -3.

If we solve the system:

$$\left(\begin{bmatrix} 1 & 1 & -1 \\ 2 & 3 & -4 \\ 4 & 1 & -4 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

we find that $x_1 = x_2 = x_3$. So we can say $x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Finding the other eigenvectors is left as an exercise.

2 Triangular Matrices

A matrix $T \in \mathbb{R}^{n \times n}$ is triangular if all values on one side of the diagonal are 0:

$$\begin{bmatrix} T_{11} & T_{12} & \dots & T_{1(n-1)} & T_{1n} \\ 0 & T_{22} & \dots & T_{2(n-1)} & T_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & T_{(n-1)(n-1)} & T_{(n-1)n} \\ 0 & 0 & \dots & 0 & T_{nn} \end{bmatrix} \text{ or } \begin{bmatrix} T_{11} & 0 & \dots & 0 & 0 \\ T_{11} & T_{22} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ T_{(n-1)1} & T_{(n-1)2} & \dots & T_{(n-1)(n-1)} & 0 \\ T_{n1} & T_{n2} & \dots & T_{n(n-1)} & T_{nn} \end{bmatrix}.$$

It turns out triangular matrices have some special properties:

- $|T| = \sum_{i=1}^{n} T_{ii}$, and
- $\langle \lambda_1, \lambda_2, ..., \lambda_n \rangle = \langle T_{11}, T_{22}, ...T_{nn} \rangle$.

Example. Find the eigenelements of $D = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$.

$$\begin{aligned} \begin{vmatrix} -\lambda & 1 & 2 & 1 \\ 0 & 1-\lambda & 0 & 2 \\ 0 & 0 & 1-\lambda & 1 \\ 0 & 0 & 0 & 2-\lambda \end{vmatrix} &= 0 \\ -\lambda \begin{vmatrix} 1-\lambda & 0 & 2 \\ 0 & 1-\lambda & 1 \\ 0 & 0 & 2-\lambda \end{vmatrix} &= 0 \\ (-\lambda)(1-\lambda) \begin{vmatrix} 1-\lambda & 1 \\ 0 & 2-\lambda \end{vmatrix} &= 0 \\ (-\lambda)(1-\lambda)^2(2-\lambda) &= 0 \end{aligned}$$

We know λs are $0, 1, 2$ (since they are the distinct values).

We can solve for each λ:

$\lambda = 0$:

$$(D - 0I_4)x = 0 \rightarrow \left(\begin{array}{ccccccc} x_2 & + & 2x_3 & + & x_4 & = & 0 \\ x_2 & + & & + & 2x_4 & = & 0 \\ & & x_3 & & & = & 0 \\ & & & & 2x_4 & = & 0 \end{array} \right) \rightarrow x = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Notice that x_1 is an arbitrary value.

$\lambda = 2$:

$$(D - 2I_4)x = 0 \rightarrow \left(\begin{array}{ccccccccc} -2x_1 & + & x_2 & + & 2x_3 & + & x_4 & = & 0 \\ & & -x_2 & + & & + & 2x_4 & = & 0 \\ & & & & -x_3 & & & = & 0 \\ & & & & & & 0 & = & 0 \end{array} \right) \rightarrow x = \begin{bmatrix} \frac{3}{2}x_4 \\ 2x_4 \\ 0 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 0 \\ 2 \end{bmatrix}$$

Notice that x_4 is an arbitrary value, but chosen to make the vector integers.

$\lambda = 1$:

$$(D - I_4)x = 0 \rightarrow \left(\begin{array}{ccccccccc} -x_1 & + & x_2 & + & 2x_3 & + & x_4 & = & 0 \\ & & & & & & 2x_4 & = & 0 \\ & & & & 0 & & & = & 0 \\ & & & & & & x_4 & = & 0 \end{array} \right) \rightarrow$$

$$\left(\begin{array}{ccccccc} -x_1 & + & x_2 & + & 2x_3 & = & 0 \\ & & & & x_4 & = & 0 \end{array} \right) \rightarrow x = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

(Note these are *bases* of the space defined by the system above.)

Symmetric Matrices. If a matrix is symmetric (that is $A = A^T$):

- all eigenvalues are real, and
- eigenvectors ore orthonormal.

3 Eigendecomposition and Diagonalization

If we create a new matrix $X \in \mathbb{R}^{n \times n}$ where each *column* is one of the eigenvectors of A, then create $\Lambda \in \mathbb{R}^{n \times n}$ which contains the eigenvalues on the diagonal (i.e. $\Lambda = \langle \lambda_1, \lambda_2, ..., \lambda_n \rangle I_n$ it turns out that

$$AX = X\Lambda$$

because it satisfies all of the eigenelement sets simultaneously.

If the eigenvectors are linearly independent then X is invertable. If X is invertable, then

$$A = X\Lambda X^{-1}.$$

We call a A ***diagonalizable*** if it can be rewritten this way. This is sometimes also called an ***eigendecomposition***.

3.1 Singlar Value Decomposition

While we're only going to briefly discuss it in this class the form

$$A = X\Lambda X^{-1}$$

is similar to whats known as the ***Singular Value Decomposition*** of a rectangular matrix $R \in \mathbb{R}^{n \times m}$ as

$$R = USV$$

(sometimes Σ is used in place of S but this is a reserved character in this class) where $U \in \mathbb{R}^{n \times n}$, $V \in \mathbb{R}^{m \times m}$, and $S \in \left\{\widehat{diag}_{n \times m}(x) \mid x \in \mathbb{R}^{min(n,m)}\right\} \subset \mathbb{R}^{n \times m}$. Here the

$$\widehat{diag}_{n \times m}\left(\langle x_1, x_2, ...x_m\rangle\right) = \begin{bmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x_m \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \text{ or } \begin{bmatrix} x_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & x_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & & \\ 0 & 0 & \dots & x_m & 0 & \dots \end{bmatrix}.$$

It produces the left matrix when $n > m$, and the right when m is larger; that is it creates a diagonal matrix padded with 0 rows/columns to make it the correct size.

Lets assume you're in the situation where you have n users who reviewed m movies and those ratings are in matrix $E \in \mathbb{R}^{n \times m}$. If you can find the SVD of the matrix such that $E = FGH$ such that $F \in \mathbb{R}^{n \times m}$, $H \in \mathbb{R}^{m \times m}$ and G is a diagonal m-dimension square matrix each of these represents something about your set. F and H show commonalities between users or movies respectively, and G is a connection matrix.

The details about finding this are beyond the scope of this course.

Useful References

Isaak and Monougian, "Basic Concepts of Linear Algebra". §7
Wilder, "10-606-f23:Lecture 6" GitHub repository, `https://github.com/bwilder0/10606-f23/blob/main/files/notes_vectorspace.pdf`
Kolter, "Linear Algebra Review and Reference", `https://www.cs.cmu.edu/~zkolter/course/15-884/linalg-review.pdf` §3.12