

Linear Algebra Primer: Part 7- Introduction to Eigentheory

Michael Levet

July 27, 2013

This tutorial will introduce the basics of eigentheory, which is the study of eigenvalues, eigenvectors, and the characteristic polynomial. Eigentheory is renowned for its usefulness in topics such as dynamical systems. This tutorial will introduce solving for eigenvalues and eigenvectors, describing their multiplicities, and diagonalizing a matrix.

Eigenvalues and Eigenvectors

Only square matrices are defined to have eigenvalues and eigenvectors. In order to solve for the eigenvalues, it is first necessary to derive the characteristic polynomial: $f(\lambda) = \det(A - \lambda I)$. The eigenvalues are the roots of the characteristic polynomial. Their associated eigenvectors satisfy the equation $A\vec{v} = \lambda\vec{v}$, where $\vec{v} \neq \vec{0}$.

Let's look at an example. Consider the matrix $A = \begin{bmatrix} 3 & 0 & 1 \\ 1 & 4 & 5 \\ 2 & 2 & 2 \end{bmatrix}$.

Its characteristic polynomial is derived from $\det(A - \lambda I)$. So $f(\lambda) = -\lambda^3 + 9\lambda^2 - 14\lambda - 12$. The eigenvalues are $\lambda = 3 + \sqrt{13}, 3 - \sqrt{13}, 3$.

Now let's solve for the eigenvectors. To do this, solve for: $(A - \lambda I)\vec{v} = \vec{0}$. Note that this form is equivalent to solving $A\vec{v} = \lambda\vec{v}$. Plug in an eigenvalue for λ and solve for the vector \vec{v} that satisfies the above equation. This vector is the associated eigenvector for the given eigenvalue.

Doing this yields the following eigenpairs (eigenvalue, eigenvector):

- $\lambda = 3 \pm \sqrt{13} : \left(\frac{\sqrt{13}}{13}, \frac{1}{26} * (13 \pm 11\sqrt{13}), 1\right)$
- $\lambda = 3 : (1, 1, 0)$

The number of times an eigenvalue occurs is called its algebraic multiplicity. The number of eigenvectors associated with each eigenvalue is called the geometric multiplicity of the eigenvalue. So in the example, each eigenvalue has an algebraic and geometric multiplicity of 1.

Deriving eigenpairs for linear transformations consists simply of deriving a transformation matrix and solving for the eigenpairs as described above. Just as with matrices, if \vec{v} is an eigenvector for a linear transformation T , then $T(\vec{v}) = \lambda\vec{v}$ and $T^n(\vec{v}) = \lambda^n\vec{v}$. So λ^n is an eigenvalue for T^n . This property can be shown by induction on $n \in \mathbb{Z}$:

Consider the base case of $n = 1$. Here, $T(\vec{v}) = \lambda\vec{v}$ by definition.

It is assumed that $T^n(\vec{v}) = \lambda^n\vec{v}$ from $n = 1$ up to an arbitrary $k \in \mathbb{Z}$. The $k+1$ case will now be proven.

Consider $T^{k+1}(\vec{v}) = T^k(T(\vec{v})) = T^k(\lambda\vec{v})$ by the inductive hypothesis. By linearity, $T^k(\lambda\vec{v}) = \lambda T^k(\vec{v})$. By the inductive hypothesis, $T^k(\vec{v}) = \lambda^k\vec{v}$, so $\lambda(\lambda^k\vec{v}) = \lambda^{k+1}\vec{v}$ by rules of exponents.

Thus, it has been assumed that $T^n(\vec{v}) = \lambda^n \vec{v}$ from $n = 1$ up to an arbitrary integer k , and proven true for the $k + 1$ case. Therefore, $T^n(\vec{v}) = \lambda^n \vec{v}$ holds true for all $n \geq 1 \in \mathbb{Z}$. Note that λ^n is an eigenvalue of T^n and \vec{v} is the associated eigenvector.

Now let's interpret eigenvalues and eigenvectors. Each eigenvalue is a solution to the equation $\det(A - \lambda I) = 0$. By the determinant test for linear independence, each eigenvalue renders a matrix linearly dependent. Similarly, an eigenvector is in the null space for $A - \lambda I$ when λ takes on the value of a specific eigenvalue. So $(1, 1, 0)$ is in the null space of $A - 3I$ in the example above.

Further more, if the algebraic and geometric multiplicities are equal for each distinct eigenvalue, then the column vectors of the matrix are linearly independent. Since the characteristic polynomial is of degree n , it follows that there are n eigenvalues. If two eigenvalues have the same eigenvector, it follows that there exist two linearly dependent columns in the matrix. Looking at it differently, Eigenspace is a subspace of the vector space represented by the matrix. Each Eigenspace has a dimension represented by the geometric multiplicity of the corresponding eigenvalue. Thus, if each eigenvalue has a distinct eigenvector, then the set of all eigenvectors form a basis for the vector space represented by the matrix.

So if there are n eigenvalues for a matrix: $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$, with the algebraic multiplicity for a given λ_i equivalent to its geometric multiplicity. Thus, the Eigenspace associated with each λ_i has a basis β_i consisting exactly of the eigenvectors associated with the eigenvalue λ_i . Thus, a basis for the vector space associated with the original matrix is $\beta = \cup_i^n \beta_i$.

Diagonalization

Using Eigenvectors as bases provides applications for diagonalization. Essentially, diagonalization allows a square matrix A to be broken up based on the eigenvalues and eigenvectors. It is in the form: $A = QDQ^{-1} = Q^{-1}DQ$, where D is the matrix whose diagonal entries are the eigenvalues of A , and Q is the matrix whose column vectors are the eigenvectors corresponding to the eigenvalues. So if $D_{11} = \lambda_1$, then the eigenvector associated with λ_1 is the first column vector in Q .

In order to diagonalize a matrix, two conditions must be met. The first condition states that the characteristic polynomial must split over the Field. This means that the characteristic polynomial can be written as a product of linear terms in the form $f(\lambda) = c(\lambda - a_1)(\lambda - a_2)\dots(\lambda - a_n)$, where $c, a_1, a_2, \dots, a_n \in \mathbb{F}$.

Consider the characteristic equation $f(\lambda) = \lambda^2 + 1$. This does not split over the real numbers as it cannot be factored; however, it does split over the complex numbers. Consider $\lambda^2 + 1 = (\lambda + i)(\lambda - i)$.

The second condition for diagonalization states that each eigenvalue must have the same algebraic and geometric multiplicity. This point was addressed to a certain extent in the previous section. If there exists an eigenvalue with a geometric multiplicity smaller than the algebraic multiplicity, then the matrix Q will not be a square matrix, as there will be fewer than n eigenvectors (given that the original matrix is $n \times n$). Thus, Q is not invertible.

The equation $A = Q^{-1}DQ$ can be thought of as a change of basis formula as well. If A is written in terms of a basis β , then Q^{-1} is a matrix that changes the basis from β to α . The matrix D is in terms of the basis α . Finally, the matrix Q changes the basis back from α to β . Often times, dealing with a different basis can provide insights or make it easier to solve certain problems.

Let's look at a couple examples of diagonalizing matrices. Consider the matrix $A = \begin{bmatrix} 6 & 2 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 8 \end{bmatrix}$.

The characteristic polynomial for A is $f(\lambda) = (\lambda - 8)(\lambda - 6)$, so it splits over the real numbers. The

eigenvector associated with $\lambda = 8$ is $(0, 0, 1)$, so the algebraic and geometric multiplicities match for $\lambda = 8$. The eigenvector associated with $\lambda = 6$ is $(1, 0, 0)$, so the algebraic multiplicity of $\lambda = 6$ is larger than its geometric multiplicity. Thus, A is not diagonalizable.

Consider a second example for the matrix $B = \begin{bmatrix} 0 & -4 \\ 2 & 6 \end{bmatrix}$.

The characteristic polynomial for B is $f(\lambda) = (\lambda - 4)(\lambda - 2)$ with the eigenvectors $(-1, 1)$ for $\lambda = 4$, and $(-2, 1)$ for $\lambda = 2$. Thus, B is diagonalizable, as its characteristic polynomial splits over the reals; and for each eigenvalue λ_i , its algebraic and geometric multiplicities are the same. So B can be written in the form $Q^{-1}DQ$ or QDQ^{-1} .

Let $Q = \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$ and $Q^{-1} = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}$.