Lecture 8: Eigenvalues, eigenvectors and diagonalization

Francesco Preta
07/30/2020

Definitions: eigenvalues and eigenvectors

Let A be a $n \times n$ matrix. We say that λ is an eigenvalue of A with eigenvector $\mathbf{v} \neq \mathbf{0}$ if

$$A\mathbf{v} = \lambda\mathbf{v}$$

Example:

$$A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Eigenspace of an eigenvalue

In general, an eigenvalue of a matrix can have more than one eigenvector.

Definition

Let A be a matrix and λ be one of its eigenvalues. The space spanned by all the eigenvectors for λ is called the **eigenspace** of λ, denoted by $E(\lambda)$. Its dimension (corresponding to the number of linearly independent eigenvectors for λ) is called the **geometric multiplicity** of λ.

Example:

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$$

0 as an eigenvalue

Notice that $\lambda = 0$ can be an eigenvalue of a matrix A, in which case the associated eigenspace is merely the Kernel of A.

The characteristic equation

We have defined eigenvalues and eigenvectors but we need to develop an algorithm to find them. Notice that for a given matrix A, we need to find both λ and $\mathbf{v} \neq \mathbf{0}$ such that $A\mathbf{v} = \lambda\mathbf{v}$. This is equivalent to solving the equation

$$(A - \lambda I)\mathbf{v} = \mathbf{0}$$

which corresponds to finding the elements of Ker$(A - \lambda I)$.

The characteristic equation (continued)

λ is also an unknown in this equation, so that we first need to find the λ for which $A - \lambda I$ has non-trivial kernel. This happens when

$$\det(A - \lambda I) = 0$$

which is called the **characteristic equation** for a matrix A. Its solutions determine the eigenvalues of A.

Example:

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

Factoring the characteristic polynomial

In general, the characteristic equation for a $n \times n$ matrix A is a polynomial of degree n. That means that it admits n (possibly complex) solutions, considering possible multiplicity. That is, we can factor it in the following way:

$$\det(A - \lambda I) = \prod_{i=1}^{k} (\lambda - \lambda_i)^{m_i}$$

so that $k \leq n, m_i \geq 1$ and $\sum_{i=1}^{k} m_i = n$. The value m_i is called the **algebraic multiplicity** of a solution λ_i.

How to find eigenvectors:

For each eigenvalue λ_i for $i = 1,...,k$ that we have found, consider the homogeneous system

$$(A - \lambda_i I)\mathbf{v} = \mathbf{0}$$

and solve it for $\mathbf{v}$. Since $\det(A - \lambda_i I) = 0$, the system will have infinitely many solutions (and not only the trivial one, $\mathbf{v} = \mathbf{0}$). Therefore the set of solutions will be a vector space spanned by a minimal set of eigenvectors. The dimension of this space will be found by looking at the solution set of $(A - \lambda_i I)\mathbf{v} = \mathbf{0}$.

In our previous example:

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}, \lambda_1 = -1, \lambda_2 = 8$$

The eigenspace is a vector subspace of $\mathbb{R}^n$

Theorem

Let λ be an eigenvalue for a square matrix $A \in \mathbb{R}^{n \times n}$. Then the eigenspace $E(\lambda)$ is a vector subspace of $\mathbb{R}^n$.

Multiplicities

We have defined the **algebraic multiplicity**, which represents the number of times a given eigenvalues solves the characteristic equation, and the **geometric multiplicity**, which represents the dimension of the corresponding eigenspace.

In general, the geometric multiplicity is less than or equal to the algebraic multiplicity. When they coincide on every eigenvector, we can find a very useful decomposition of a square matrix: its **diagonalization**.

Example:

$$A = \begin{bmatrix} 5 & 2 & 4 \\ 0 & 4 & 1 \\ 0 & 0 & 5 \end{bmatrix}$$

Eigenvectors of different eigenvalues are linearly independent

Theorem

Let $\mathbf{v}_1 \in E(\lambda_1)$ and $\mathbf{v}_2 \in E(\lambda_2)$ be two nontrivial eigenvectors for a square matrix $A \in \mathbb{R}^{n \times n}$, with λ_1 and λ_2 two different eigenvalues of A. Then $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent.

Proof:

Diagonalization

If there are n eigenvectors of A, $\mathbf{v}_1, \ldots, \mathbf{v}_n$, that are linearly independent, then we can apply a change of basis, from the canonical basis to $\mathbf{v}_1, \ldots, \mathbf{v}_n$ to observe the action of A on these eigenspaces.

Diagonalization (continued)

In particular, for any $\mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$, for $\mathbf{v}_i \in E(\lambda_i)$, we get

$$A\mathbf{x} = \sum_{i=1}^{n} \alpha_i A \mathbf{v}_i = \sum_{i=1}^{n} \alpha_i \lambda_i \mathbf{v}_i$$

This shows that in a basis made of eigenvectors, A acts by multiplying each one of them by a constant. This is the same type of action of a diagonal matrix on the canonical basis, therefore it induces a **diagonalization** of A.

When is a matrix diagonalizable?

Theorem

Let A be a square $n \times n$ matrix with n independent eigenvectors. Then there exists an invertible matrix P and a diagonal matrix Λ such that $A = P\Lambda P^{-1}$. For a given ordering of n independent eigenvectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, the matrix P is a matrix having i-th column equal to $\mathbf{v}_i$ and the matrix Λ has on the i-th element of the diagonal the eigenvalue λ_i corresponding to $\mathbf{v}_i$.

Example:

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

Interpretation of the diagonalization:

The meaning of this decomposition can be seen as such: P^{-1} is the change of basis matrix from the canonical basis $\{\mathbf{e}_i\}_{i=1}^n$ to the basis of eigenvectors $\{\mathbf{v}_i\}_{i=1}^n$. That means that if $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i = \sum_{i=1}^n \alpha_i \mathbf{v}_i$, then

$$P\Lambda P^{-1} \sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \sum_{i=1}^{n} \alpha_i P\Lambda \mathbf{e}_i = \sum_{i=1}^{n} \alpha_i \lambda_i P \mathbf{e}_i = \sum_{i=1}^{n} \alpha_i \lambda_i \mathbf{v}_i = A \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$$

so that A and $P\Lambda P^{-1}$ coincide for every $\mathbf{x} \in \mathbb{R}^n$.

Conditions for diagonalization

Theorem

Let A be a $n \times n$ matrix. Then A is diagonalizable if and only if

for every eigenvalue λ_i of A, the algebraic multiplicity and the geometric multiplicity coincide. In particular, if A has all distinct eigenvalues, it is diagonalizable.

Example:

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

Is it diagonalizable? If yes, find a diagonalization.

How to solve diagonalization problems

1. Find the eigenvalues.
   If they are all different, the matrix is diagonalizable.
2. Find the corresponding eigenvectors.
   At this point, we would always be able to find out if the matrix is diagonalizable, depending on the multiplicities of its eigenvalues.
3. **If it is diagonalizable:** write down the matrices that compose the diagonalization.

Application of diagonalization

If we need to calculate a power of a given matrix A, say A^k, we can avoid computing $k-1$ matrix multiplications, but rather we can diagonalize it to obtain:

$$A^k = (P\Lambda P^{-1})^k = P\Lambda P^{-1} \cdot P\Lambda P^{-1} \cdot \ldots \cdot P\Lambda P^{-1} = P\Lambda^k P^{-1}$$

and notice that computing Λ^k corresponds only to raise each element on the diagonal to the k-th power.

Complex eigenvalues

Sometimes the characteristic equation of a matrix having real coefficients doesn't have solutions in $\mathbb{R}$.

Example:

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

Fundamental theorem of algebra

Let p be a polynomial of degree n. Then $p(\lambda) = 0$ admits n solutions in $\mathbb{C}$, possibly with multiplicity. In particular, if p has real coefficients and $\lambda_0 \in \mathbb{C}$ is a solution to $p(\lambda) = 0$, then its complex conjugate $\bar{\lambda}_0$ is also a solution to $p(\lambda) = 0$.

Complex solutions of the characteristic equation

In particular, the characteristic polynomial of a $n \times n$ matrix with real coefficients is always a polynomial of degree n. It will admit n solutions, up to multiplicity and could potentially admit pairs of conjugate complex solutions. In this case, diagonalization is not really effective, since the eigenvectors will also be in $\mathbb{C}$.

Example:

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \lambda_1 = i, \lambda_2 = -i$$

Another type of decomposition

In the 2×2 case we can actually consider another type of decomposition that gives us more information: for any A having a pair of complex eigenvalues, we want to find a conjugate of the form

$$C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

Conjugate rotation

For $r = \sqrt{a^2 + b^2}$, we can decompose it into

$$C = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \frac{a}{r} & -\frac{b}{r} \\ \frac{b}{r} & \frac{a}{r} \end{bmatrix}$$

and treat the first matrix as a dilatation by r along both axis and the second as a rotation by a θ such that $\cos(\theta) = \frac{a}{r}$ and $\sin(\theta) = \frac{b}{r}$.

Conjugating matrix

Then for a given $A \in \mathbb{R}^{2\times 2}$ having complex eigenvalues, let $\mathbf{v}$ be one of the two eigenvectors and take

$$S = \begin{bmatrix} \text{Re}(\mathbf{v}) & \text{I}(\mathbf{v}) \end{bmatrix}$$

where $\text{Re}(\mathbf{v})$ is the real part and $\text{I}(\mathbf{v})$ is the imaginary part of $\mathbf{v}$. Then there exists a C such that

$$A = SCS^{-1}$$

Example:

$$A = \begin{bmatrix} \frac{1}{2} & -\frac{3}{5} \\ \frac{3}{4} & \frac{11}{10} \end{bmatrix}$$

Complex eigenvalues=rotations

The decomposition $A = SCS^{-1}$ works as follows: S represents a change of basis matrix, C represents a rotation matrix (possibly with dilatation by the same amount in both directions) and S^{-1} represents a return to the original basis. To this extent, every 2×2 matrix with complex eigenvalues is, in some sense, a rotation.

Trace of a matrix

Definition

Let A be a $n \times n$ matrix, then we define the **Trace** of A to be

$$\text{Tr}(A) = \sum_{i=1}^{n} A_{i,i}$$

Spectral properties

Theorem

Let A be a $n \times n$ matrix and Λ the diagonal matrix with its eigenvalues, then

$$\text{Tr}(A) = \text{Tr}(\Lambda) = \sum_{i=1}^{n} \lambda_i$$

$$\det(A) = \det(\Lambda) = \prod_{i=1}^{n} \lambda_i$$

$\text{Rank}(A) = \text{Rank}(\Lambda) =$ number of nonzero eigenvalues of A

Spectral properties (continued)

The theorem can be used both ways: if we already have an eigendecomposition, we can calculate Trace, Rank and Determinant of a matrix through its eigenvalues. On the other hand, we can calculate eigenvalues of 2×2 matrix A just by considering Tr(A) and $\det(A)$.

Example:

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

Symmetric matrices

Theorem

Let A be a symmetric $n \times n$ matrix. Then A admits a diagonalization by n orthonormal eigenvectors.

Proof of orthogonality for different eigenspaces of a symmetric matrix

Example:

$$A = \begin{bmatrix} 1 & -2 & 2 \\ -2 & -2 & 4 \\ 2 & 4 & -2 \end{bmatrix}$$

Diagonalization of a symmetric matrix

Notice that the diagonalization is easier to compute for symmetric matrices since if U is the matrix having the i-th unitary eigenvector as i-th column, then $U^{-1} = U^T$, so that $A = U\Lambda U^T$.