DIAGONALIZATION PART 01

JOSÉ MALAGÓN-LÓPEZ

Motivated by the fact that it is relatively easier to answer questions about diagonal matrices, in this lecture we deal with the problem of determining which matrices are "similar" to a diagonal matrix. We will see that this problem is closely related to the problem of finding vectors in $\mathbb{R}^n$ for which $A\vec{v} = \lambda\vec{v}$, for some λ.

1. Diagonalizable Matrices

Definition 1. Let A and B two $n \times n$ matrices. We say that A and B are **similar** if there is an invertible matrix Q such that

$$A = QBQ^{-1}.$$

Theorem 2. *If A and B are two similar $n \times n$ matrices, then*

$$\det(A) = \det(B).$$

Proof. If A and B are similar, by definition there is an invertible matrix Q such that $A = QBQ^{-1}$. Then

$$\begin{aligned}
\det(A) &= \det\left(QBQ^{-1}\right) \\
&= \det(Q)\det(B)\det\left(Q^{-1}\right) \\
&= \det(Q)\det\left(Q^{-1}\right)\det(B) \\
&= \det\left(QQ^{-1}\right)\det(B) \\
&= \det(I_n)\det(B) \\
&= \det(B).
\end{aligned}$$

Q.E.D.

Definition 3. A square matrix is called **diagonalizable** if it is similar to a diagonal matrix.

The problem of determining if a square matrix is diagonalizable, and if it is, how to obtain the expression of the form $A = QDQ^{-1}$, where D is a diagonal matrix, can be solved using the remark below.

Remark 4. By definition, an 2×2 matrix A

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is diagonalizable if and only if there is an invertible matrix Q and a diagonal matrix D such that $AQ = QD$. Assume

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

with λ_1, λ_2 real numbers not necessarily different.

To say that Q is an invertible matrix is equivalent to saying that

$$\text{Col}(Q) = \mathbb{R}^2.$$

Thus, the columns of Q consist of 2 linearly independent vectors in $\mathbb{R}^2$. Assume

$$Q = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}.$$

Now, with these notations, from one side

$$AQ = \begin{pmatrix} A \begin{pmatrix} q_{11} \\ q_{21} \end{pmatrix} & A \begin{pmatrix} q_{12} \\ q_{22} \end{pmatrix} \end{pmatrix}.$$

From another side

$$QD = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 q_{11} & \lambda_2 q_{12} \\ \lambda_1 q_{21} & \lambda_2 q_{22} \end{pmatrix}.$$

In other words, $AQ = QD$ is equivalent to having

$$A \begin{pmatrix} q_{11} \\ q_{21} \end{pmatrix} = \lambda_1 \begin{pmatrix} q_{11} \\ q_{21} \end{pmatrix}, \qquad A \begin{pmatrix} q_{12} \\ q_{22} \end{pmatrix} = \lambda_2 \begin{pmatrix} q_{12} \\ q_{22} \end{pmatrix}.$$

The previous example illustrate the general situation.

Theorem 5. *An $n \times n$ matrix A is diagonalizable if and only if there is a basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ for $\mathbb{R}^n$ such that for all j*

$$A\vec{v}_j = \lambda_j \vec{v}_j,$$

for some number λ_j. In this case, $A = QDQ^{-1}$, where

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

and Q is the matrix whose j-th column is the vector $\vec{v}_j$.

2. Eigenvalues and Eigenvectors

Definition 6. Let A be an $n \times n$ matrix. A non-zero vector $\vec{v}$ in $\mathbb{R}^n$ is an **eigenvector** if $A\vec{v} = \lambda\vec{v}$, for some scalar λ.

The scalar λ is called the **eigenvalue** of A corresponding to the eigenvector $\vec{v}$.

Sometimes the words proper vector (value) or characteristic vector (value) are used instead of eigenvector (eigenvalue).

Example 7. Since

$$\begin{pmatrix} 2 & 4 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 28 \\ 35 \end{pmatrix} = 7 \begin{pmatrix} 4 \\ 5 \end{pmatrix},$$

we have that $\begin{pmatrix} 4 \\ 5 \end{pmatrix}$ is an eigenvector of $\begin{pmatrix} 2 & 4 \\ 5 & 3 \end{pmatrix}$ with 7 as its eigenvalue.

Example 8. Since

$$\begin{pmatrix} 1 & 9 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 10 \\ 2 \end{pmatrix} \neq \lambda \begin{pmatrix} 3 \\ 1 \end{pmatrix},$$

for any λ, we have that $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is NOT an eigenvector of $\begin{pmatrix} 1 & 9 \\ 1 & 1 \end{pmatrix}$.

Remark 9. (1) Eigenvectors must be non-zero, but the eigenvalues can be zero.

(2) Different vectors can have same eigenvalues. More precisely, if $A\vec{v} = \lambda\vec{v}$ then for any non-zero scalar α, the vector $\alpha\vec{v}$ is also an eigenvector of A:

$$A\,(\alpha\vec{v}) = \alpha\,(A\vec{v}) = \alpha(\lambda\vec{v}) = \lambda\,(\alpha\vec{v})\,.$$

(3) If $A = QDQ^{-1}$ then $A^l = QD^lQ^{-1}$ for any integer $l > 0$.

First we will find the eigenvalues of a matrix and then we will find their corresponding eigenvectors.

2.1. **Finding the Eigenvalues of a Matrix.** The first theorem gives us a criterion to determine when a number is an eigenvalue of a matrix.

Theorem 10. *Let A be an $n \times n$ matrix. A scalar λ is an eigenvalue of A if and only if*

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

Proof. The equality $A\vec{x} = \lambda\vec{x}$ is equivalent to having

$$\vec{0} = A\vec{x} - \lambda\vec{x} = A\vec{x} - \lambda(I_n\vec{x}) = (A - \lambda I_n)\,\vec{x}.$$

Thus, λ is an eigenvalue of A if and only if there is a vector $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x} = \lambda\vec{x}$, which is equivalent to the system

$$(A - \lambda I_n)\,\vec{x} = \vec{0}.$$

Since eigenvectors can not be zero-vectors, the system above must have non-trivial solutions. This is equivalent to having

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

Q.E.D.

The use of the following definition is clear from the previous theorem.

Definition 11. Let A be an $n \times n$ matrix. The polynomial

$$P(t) = \det\,(A - t\,I)$$

is called the **characteristic polynomial** of A.

The following theorem not only determines if a matrix has eigenvalues, but if they exist, it will give us explicitly such values.

Theorem 12. *Let A be an $n \times n$ matrix.*

(1) *The characteristic polynomial* $\det(A - t\,I)$ *has degree* n *and leading coefficient* $(-1)^n$.

(2) *A scalar is an eigenvalue of A if and only if $P(\lambda) = 0$.*

(3) *A has n distinct eigenvalues, possibly complex numbers.*

Example 13. Consider the matrix $A = \begin{pmatrix} 2 & 4 \\ 5 & 3 \end{pmatrix}$. Its characteristic polynomial is

$$\begin{vmatrix} 2-t & 4 \\ 5 & 3-t \end{vmatrix} = (t-7)(t+2)$$

We conclude that the eigenvalues of A are $\{7, -2\}$.

Example 14. Consider the matrix $A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$. Its characteristic polynomial is

$$\det(A - tI) = \begin{vmatrix} 1-t & 2 \\ -1 & 1-t \end{vmatrix} = t^2 - 2t + 3$$

We conclude that the eigenvalues of A are $\left\{1 + i\sqrt{2}, 1 - i\sqrt{2}\right\}$.

Example 15. Consider the matrix $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$. Its characteristic polynomial is

$$\det(A - tI) = \begin{vmatrix} 3-t & 1 & 0 \\ 0 & 3-t & 0 \\ 0 & 0 & 4-t \end{vmatrix} = -(t-4)(t-3)^2$$

We conclude that the eigenvalues of A are $\{3, 4\}$.

2.2. **Multiplicities of Eigenvalues.**

Definition 16. The **(algebraic) multiplicity** of an eigenvalue λ is the largest number m for which $(t-\lambda)^m$ is a factor of $\det(A - tI)$.

Example 17. Consider the matrix $A = \begin{pmatrix} 2 & 4 \\ 5 & 3 \end{pmatrix}$. We saw that

$$\det(A - tI) = (t-7)(t+2).$$

We conclude that $\lambda = 7$ is an eigenvalue of A of multiplicity one, and $\lambda = -2$ is an eigenvalue of A of multiplicity one.

Example 18. Consider the matrix $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$. We saw that

$$\det(A - tI) = (3-t)(3-t)(4-t) = -(t-4)(t-3)^2$$

We conclude that $\lambda = 4$ is an eigenvalue of A of multiplicity one, and $\lambda = 3$ is an eigenvalue of A of multiplicity 2.

2.3. **Finding the Eigenvectors of a Matrix.**

The following theorem gives us a criterion to determine when a vector is an eigenvector of a matrix.

Theorem 19. *Let A be an $n \times n$ matrix. Let λ be an eigenvalue of A. A vector $\vec{v} \neq \vec{0}$ is an eigenvector of A with eigenvalue λ if and only if $\vec{v}$ is in* $\mathrm{Null}(A - \lambda I_n)$.

Definition 20. Let A be an $n \times n$ matrix. Let λ be an eigenvalue of A. The subspace

$$\mathrm{E}_\lambda = \mathrm{Null}(A - \lambda I)$$

is called the **eigenspace of A corresponding to** λ.

Remark 21. Consider a matrix A and an eigenvalue λ of A. Then

(1) **Every** vector in E_λ is an eigenvector of A with eigenvalue λ.
(2) $E_\lambda \neq \{\vec{0}\}$ for any eigenvalue λ.

Theorem 22. *Let A be an $n \times n$ matrix. Let λ be an eigenvalue of A of multiplicity m. Then*

$$1 \leq \dim(\mathrm{E}_\lambda) \leq m.$$

3. Criterion for Diagonalizing a Matrix

Now we state the theorem that will give us a test for determining if a matrix is diagonalizable.

Theorem 23. *Let A be an $n \times n$ matrix. Then*

(1) *A is diagonalizable if and only if*

$$\det(A - tI) = (-1)^n (t - \lambda_1) \cdots (t - \lambda_n)$$

and $\dim(\mathrm{E}_{\lambda_i}) = m_i$ for all i, where m_i is the multiplicity of λ_i.

(2) *If A is diagonalizable and B_i is an ordered basis of E_{λ_i}, then*

$$B = B_1 \cup \cdots \cup B_n$$

is a basis of $\mathbb{R}^n$ consisting only of eigenvectors.

(3) *Let D be the diagonal matrix whose entries along the diagonal are the eigenvalues of A, where λ_i appears m_i times, Let Q be the matrix whose columns are the vectors in the basis B. Then*

$$A = QDQ^{-1}.$$

E-mail address: jmalagon@uottawa.ca