

블로그 메뉴

홈  
미디어로그  
태그  
방명록

## 검색결과 리스트

### [선형대수학 #3] 고유값과 고유벡터 (eigenvalue & eigenvector)

수학 이야기 2013. 10. 16. 14:35

선형대수학에서 고유값(eigenvalue)과 고유벡터(eigenvector)가 중요하다고 하는데 왜 그런 것인지 개인적으로도 궁금합니다. 고유값, 고유벡터에 대한 수학적 정의 말고 이런 것들이 왜 나왔고 그 본질이 무엇인지에 대한 직관이 있으면 좋을텐데요..

아직은 딱히 이것 때문이라고 결론지을 수는 없지만 고유값, 고유벡터 그 자체의 활용보다는 SVD(특이값분해), Pseudo-Inverse, 선형연립방정식의 풀이, PCA(주성분분석) 등의 주요 응용이 eigenvalue, eigenvector를 그 밑바탕에 깔고 있기 때문은 아닌가 생각하고 있습니다.

이 글에서는 먼저 고유값, 고유벡터 자체에 대한 개념만 살펴보고 SVD, 선형연립방정식, PCA 등에 대해서는 별도 글로 자세하게 다룰 예정입니다.

#### 1. 고유값, 고유벡터란?

고유값(eigenvalue), 고유벡터(eigenvector)에 대한 수학적 정의는 비교적 간단하다.

행렬 A를 선형변환으로 봤을 때, **선형변환 A에 의한 변환 결과가 자기 자신의 상수배가 되는 0이 아닌 벡터**를 고유벡터(eigenvector)라 하고 이 상수배 값을 고유값(eigenvalue)라 한다.

즉,  $n \times n$  정방행렬(**고유값, 고유벡터는 정방행렬에 대해서만 정의된다**) A에 대해  $A\mathbf{v} = \lambda\mathbf{v}$ 를 만족하는 0이 아닌 열벡터  $\mathbf{v}$ 를 고유벡터, 상수  $\lambda$ 를 고유값이라 정의한다.

$$A\mathbf{v} = \lambda\mathbf{v} \quad \text{---(1)}$$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad \text{--- (2)}$$

좀더 정확한 용어로는  $\lambda$ 는 '행렬 A의 고유값',  $\mathbf{v}$ 는 '행렬 A의  $\lambda$ 에 대한 고유벡터'이다.

즉, 고유값과 고유벡터는 행렬에 따라 정의되는 값으로서 어떤 행렬은 이러한 고유값-고유벡터가 아예 존재하지 않을수도 있고 어떤 행렬은 하나만 존재하거나 또는 최대  $n$ 개까지 존재할 수 있다.

## 2. 기하학적 의미

기하학적으로 봤을 때, 행렬(선형변환) A의 고유벡터는 선형변환 A에 의해 방향은 보존되고 스케일(scale)만 변화되는 방향 벡터를 나타내고 고유값은 그 고유벡터의 변화되는 스케일 정도를 나타내는 값이다.



예를 들어 지구의 자전운동과 같이 3차원 회전변환을 생각했을 때, 이 회전변환에 의해 변하지 않는 고유벡터는 회전축 벡터이고 그 고유값은 1이 될 것이다.

## 3. 고유값분해를 이용한 대각화 - eigendecomposition

고유값, 고유벡터는 정방행렬의 대각화와 밀접한 관련이 있다 (eigendecomposition은 정방행렬에 대해서만 가능함)

먼저 대각행렬과의 행렬곱에 대해 살펴보면, 대각행렬을 뒤에 곱하면 행렬의 열벡터들이 대각원소의 크기만큼 상수배가 된다(앞에 곱하면 행벡터들이 상수배가 된다). 예를 들어, 3 x 3 행렬의 경우를 보면 다음과 같다.

$$\begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = \begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} & \lambda_3 v_{13} \\ \lambda_1 v_{21} & \lambda_2 v_{22} & \lambda_3 v_{23} \\ \lambda_1 v_{31} & \lambda_2 v_{32} & \lambda_3 v_{33} \end{pmatrix} \quad \text{--- (3)}$$

행렬 A의 고유값, 고유벡터들을  $\lambda_i, \mathbf{v}_i, i = 1, 2, \dots, n$ 이라 하자.

$$\begin{aligned} A\mathbf{v}_1 &= \lambda_1 \mathbf{v}_1 \\ A\mathbf{v}_2 &= \lambda_2 \mathbf{v}_2 \\ &\vdots \\ A\mathbf{v}_n &= \lambda_n \mathbf{v}_n \quad \text{--- (4)} \end{aligned}$$

이제 식 (4)를 한꺼번에 표현하여 정리하면

$$A[\mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_n] = [\lambda_1 \mathbf{v}_1 \lambda_2 \mathbf{v}_2 \cdots \lambda_n \mathbf{v}_n]$$

$$= [\mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_n] \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_n \end{bmatrix} \quad \text{--- (5)}$$

가 성립함을 알 수 있다.

즉, 행렬  $A$ 의 고유벡터들을 열벡터로 하는 행렬을  $P$ , 고유값들을 대각원소로 하는 대각행렬을  $\Lambda$ 라 하면 다음 식이 성립한다.

$$AP = P\Lambda \quad \text{--- (6)}$$

$$\text{즉, } A = P\Lambda P^{-1} \quad \text{--- (7)}$$

이와같이 행렬  $A$ 는 자신의 고유벡터들을 열벡터로 하는 행렬과 고유값을 대각원소로 하는 행렬의 곱으로 대각화 분해가 가능한데 이러한 대각화 분해를 eigendecomposition이라고 한다.

한 예로,  $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$ 인 경우  $A$ 는 다음과 같이 대각화가 가능하다.

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}^{-1} \quad \text{--- (8)}$$

모든 정방행렬이 이런 방식의 eigendecomposition이 가능한 것은 아니지만 대각화 가능한 경우는 뒤에 적기로 하고 일단은 대각화를 하면 어떻게 좋은지 알아보자.

행렬  $A$ 의 eigendecomposition을 알면 행렬식 값  $\det(A)$ ,  $A$ 의 거듭제곱, 역행렬, 대각합(trace), 행렬의 다항식 등을 매우 손쉽게 계산할 수 있다.

$$\begin{aligned} \det(A) &= \det(P\Lambda P^{-1}) \\ &= \det(P)\det(\Lambda)\det(P)^{-1} \\ &= \det(\Lambda) \\ &= \lambda_1 \lambda_2 \cdots \lambda_n \end{aligned} \quad \text{--- (9)}$$

$$\begin{aligned} A^k &= (P\Lambda P^{-1})^k \\ &= (P\Lambda P^{-1})(P\Lambda P^{-1}) \cdots (P\Lambda P^{-1}) \\ &= P\Lambda^k P^{-1} \\ &= P \text{diag}(\lambda_1^k, \dots, \lambda_n^k) P^{-1} \end{aligned} \quad \text{--- (10)}$$

$$\begin{aligned} A^{-1} &= (P\Lambda P^{-1})^{-1} \\ &= P\Lambda^{-1}P^{-1} \\ &= P \text{diag}(1/\lambda_1, \dots, 1/\lambda_n) P^{-1} \quad \dots (11) \end{aligned}$$

$$\begin{aligned} \text{tr}(A) &= \text{tr}(P\Lambda P^{-1}) \\ &= \text{tr}(\Lambda) \quad (\because \text{tr}(AB) = \text{tr}(BA)) \\ &= \lambda_1 + \dots + \lambda_n \quad \dots (12) \end{aligned}$$

$$\begin{aligned} f(A) &= a_0 E + a_1 A + \dots + a_n A^n \quad (f(x) = a_0 + a_1 x + \dots + a_n x^n) \\ &= a_0 P P^{-1} + a_1 P \Lambda P^{-1} + \dots + a_n P \Lambda^n P^{-1} \\ &= P(a_0 P^{-1} + a_1 \Lambda P^{-1} + \dots + a_n \Lambda^n P^{-1}) \\ &= P(a_0 E + a_1 \Lambda + \dots + a_n \Lambda^n) P^{-1} \\ &= P \text{diag}(f(\lambda_1), \dots, f(\lambda_n)) P^{-1} \quad \dots (13) \end{aligned}$$

#### 4. 고유값분해(eigendecomposition) 가능조건 - 일차독립

앞서 말했지만 모든 정방행렬이 고유값분해가 가능한 것은 아니다.  $n \times n$  정방행렬  $A$ 가 **고유값분해가 가능하려면 행렬  $A$ 가  $n$ 개의 일차독립인 고유벡터**를 가져야 한다.

선형대수학에서 말하는 일차독립(linearly independent)이란 무엇일까?

어떤 벡터들의 집합  $\{v_1, \dots, v_n\}$ 이 있을 때, 이들 벡터들 중 어느 한 벡터도 다른 벡터들의 일차결합으로 표현될 수 없으면 이 벡터들은 서로 일차독립이라고 정의한다.

벡터들의 일차결합이란  $a_1 v_1 + a_2 v_2 + \dots + a_n v_n$  ( $a_i$ 는 상수)와 같이 상수를 곱하여 합친 형태를 말한다.

예를 들어 3차원 공간의 좌표축 단위벡터들인  $v_1 = (1, 0, 0)$ ,  $v_2 = (0, 1, 0)$ ,  $v_3 = (0, 0, 1)$ 를 생각해 보면  $v_2$ ,  $v_3$ 에 어떤 상수를 곱하여 더해도  $v_1$ 이 나올수 없음은 쉽게 확인할 수 있다. 이와 같이 어떤 벡터도 다른 벡터들의 상수배 합으로 표현될 수 없으면 이 벡터들은 서로 일차독립(linearly independent)이라고 한다.

참고로  $n$ 차원 공간은 최대  $n$ 개의 일차독립인 벡터들을 가질 수 있으며  $n$ 개의 일차독립인 벡터들은 이 공간을 생성하는 기저(basis) 역할을 수행한다.

위 예에서, 3차원 공간에서 좌표축 단위벡터들의 집합  $\{v_1, v_2, v_3\}$ 는 일차독립이지만 여기에  $v_4 = (-1, 3, 4)$ 를 추가한  $\{v_1, v_2, v_3, v_4\}$ 는 일차독립이 아니다 ( $\because v_4 = -v_1 + 3v_2 + 4v_3$ ). 즉, 3차원 공간에서 가능한 일차독립인 벡터들의 개수는 최대 3개이다.

또한  $v_1, v_2, v_3$ 를 자세히 보면 이 벡터들을 이용하여 3차원 공간의 모든  $(x, y, z)$  좌표를 생성할 수 있음을 알 수 있다. 어떤 일차독립인 벡터들의 집합  $\{v_1, \dots, v_n\}$ 의 일차결합을 통해 어떤 공간의 모든 좌표를 생성할 수 있으면 이 벡터들이 이 공간의 기저(basis)라고 정의한다.

다시 원래 문제로 돌아가서,  $n$ 차 정방행렬이 고유값분해가 가능하려면  $n$ 개의 일차독립인 고유벡터가 존재해야 한다고 했는데 이게 무슨 말인지 아직은 마음에 확 와닿지는 않을 것이다. 그건 고유값, 고유벡터가 어떻게 계산되는지 그 과정

에 대한 이해가 필요하기 때문이다.

## 5. 고유값, 고유벡터의 계산

고유값과 고유벡터를 정의하는 식인 식 (1)으로 다시 돌아가서 이를  $\mathbf{v}$ 에 대해 정리해보면 다음과 같다.

$$\begin{aligned} A\mathbf{v} &= \lambda\mathbf{v} \\ A\mathbf{v} - \lambda\mathbf{v} &= \mathbf{0} \quad (\mathbf{0}: \text{영행렬}) \\ (A - \lambda E)\mathbf{v} &= \mathbf{0} \quad (E: \text{단위행렬}) \quad \text{--- (14)} \end{aligned}$$

우리가 구하고자 하는 고유값, 고유벡터는 식 (14)를 풀어서 나오는  $\lambda$ 와  $\mathbf{v}$ 이다 (단,  $\mathbf{v} \neq \mathbf{0}$ ). 그런데, 식 (14)를 잘 보면  $(A - \lambda E)$ 의 역행렬이 존재한다면  $\mathbf{v}$ 는 항상  $\mathbf{v} = (A - \lambda E)^{-1}\mathbf{0} = \mathbf{0}$  만을 해로 갖게 된다. 그런데, 고유벡터는 정의에 의해 영벡터가 아닌 벡터여야 하므로  $A - \lambda E$ 의 역행렬이 존재하지 않는 경우에만 존재할 수 있다. 따라서, 고유벡터가 존재하기 위해서는 일단은  $\det(A - \lambda E) = 0$  이어야 한다.

$$\det(A - \lambda E) = 0 \quad \text{--- (15)}$$

이 때, 식 (15)를 행렬  $A$ 의 특성방정식(characteristic equation)이라고 부르며 식 (15)를  $\lambda$ 에 대해 풀면  $A$ 의 고유값을 구할 수 있다. 고유벡터는 이렇게 구한  $\lambda$ 를 다시 식 (14)에 대입하여 계산한다.

[예]

다음과 같은 행렬  $A$ 에 대한 고유값, 고유벡터를 계산해 보자.

$$A = \begin{bmatrix} 2 & 0 & -2 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{--- (16)}$$

이 때, 행렬  $A$ 를 식 (14)에 대입하여 특성다항식을 구해보면

$$\begin{aligned} \det(A - \lambda E) &= \det\left(\begin{bmatrix} 2 & 0 & -2 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) \\ &= \begin{vmatrix} 2-\lambda & 0 & -2 \\ 1 & 1-\lambda & -2 \\ 0 & 0 & 1-\lambda \end{vmatrix} \\ &= (2-\lambda)((1-\lambda)(1-\lambda) - 0) \\ &= (2-\lambda)(1-\lambda)^2 \quad \text{--- (17)} \end{aligned}$$

이므로 특성방정식은  $(2-\lambda)(1-\lambda)^2 = 0$  이 된다.

$$(2-\lambda)(1-\lambda)^2 = 0 \quad \text{--- (18)}$$

이제 특성방정식의 해는  $\lambda = 1, 2$ 인데 잘 보면  $\lambda=2$ 는 단일근임에 비해  $\lambda=1$ 은 이중근임을 알 수 있다.  $\lambda$ 에 대응하는 고

유벡터의 개수는  $\lambda$ 가 몇중근이냐와 밀접한 관계가 있는데 단일근에 대해서는 1개, 이중근에 대해서는 최대 2개, 삼중근에 대해서는 최대 3개, ... 의 고유벡터가 존재한다.

먼저  $\lambda = 2$ 를 다시 식 (14)에 대입하여 고유벡터를 구해보면,

$$\begin{bmatrix} 2-\lambda & 0 & -2 \\ 1 & 1-\lambda & -2 \\ 0 & 0 & 1-\lambda \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & -2 \\ 1 & -1 & -2 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$-2v_z=0, v_x-v_y-2v_z=0, -v_z=0$$

$$\therefore v_x=v_y, v_z=0 \quad \text{--- (19)}$$

따라서,  $\lambda = 2$ 에 대응하는 고유벡터는  $v = [1, 1, 0]^T$ 로 잡을 수 있다.

마찬가지로,  $\lambda = 1$ 에 대해서도 고유벡터를 구해보면

$$\begin{bmatrix} 1 & 0 & -2 \\ 1 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\therefore v_x=2v_z \quad \text{--- (20)}$$

x좌표가 z좌표의 2배인 벡터들은 무수히 많은데 이들은  $[2, 0, 1], [0, 1, 0]$ 의 일차결합으로 표현할 수 있으므로  $\lambda = 1$ 에 대응하는 고유벡터를  $[2, 0, 1]$ 과  $[0, 1, 0]$ 으로 잡을 수 있다.

☞ x좌표가 z좌표의 2배인 임의의 벡터는  $[2t, s, t]$ 로 표현할 수 있다 (단,  $t, s$ 는 임의의 실수). 그런데, 이는  $[2t, s, t] = t*[2, 0, 1] + s*[0, 1, 0]$ 와 같이  $[2, 0, 1]$ 과  $[0, 1, 0]$ 의 일차결합으로 표현될 수 있다.

아마도 이 시점에서 고유값은 어떻게 구하는지 알겠는데 고유벡터는 뭔가 좀 애매하다고 느끼는 분들이 많을 것으로 예상된다. 그 이유는 어떤 행렬에 대해 고유값은 유일하게 결정되지만 고유벡터는 유일하지 않기 때문이다. 이는 다음 내용을 보면 알 수 있다.

식(1)의 양변에 상수  $c$ 를 곱해보면,

$$A(cv)=\lambda(cv) \quad \text{--- (21)}$$

가 되므로  $v$ 가  $\lambda$ 에 대한 고유벡터이면 0이 아닌 임의의 상수  $c$ 에 대해  $cv$ 도 또한  $\lambda$ 에 대한 고유벡터임을 알 수 있다.

또한,  $v_1, v_2$ 가 모두 고유값  $\lambda$ 에 대응하는 고유벡터라고 하면 임의의 상수  $c_1, c_2$ 에 대해

$$A(c_1v_1+c_2v_2)=c_1\lambda v_1+c_2\lambda v_2=\lambda(c_1v_1+c_2v_2) \quad \text{--- (22)}$$

이므로  $c_1v_1 + c_2v_2$  또한  $\lambda$ 에 대한 고유벡터가 됨을 알 수 있다.

따라서 고유벡터는 식 (19), 식 (20) 등과 같은 제약조건을 만족하는 벡터들 중에서 어느 벡터를 사용해도 무방하나 보통은 벡터의 크기를 1로 정규화한(normalized) 단위벡터를 고유벡터로 잡는 것이 일반적이다. 단, 식 (20)의 경우에는 자유도(degree of freedom)가 2이기 때문에 일차독립인 2개의 고유벡터를 잡아야만 가능한 모든 고유벡터들을 대표할 수 있다.

이제 이렇게 구한 고유값, 고유벡터에 대해 아래와 같은 행렬 대각화가 정말로 성립함을 matlab 또는 손계산을 이용하면 쉽게 확인할 수 있다 (또한 고유벡터들의 스케일을 바꾸거나 순서를 바꾸어도 대각화가 성립함도 확인할 수 있다).

$$\begin{bmatrix} 2 & 0 & -2 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \quad \text{--- (23)}$$

## 6. 대칭행렬(symmetric matrix)과 고유값분해

정방행렬들 중에서 대각원소를 중심으로 원소값들이 대칭되는 행렬 즉,  $A^T = A$  (모든  $i, j$ 에 대해  $a_{ij} = a_{ji}$ )인 행렬을 대칭행렬(symmetric matrix)이라 부른다.

그런데, 대칭행렬은 고유값 분해와 관련하여 매우 좋은 성질 2가지를 가지고 있다. **실원소(real-valued) 대칭행렬은 항상 고유값 대각화가 가능하며 더구나 직교행렬(orthogonal matrix)로 대각화가 가능하다는** 매우 좋은 성질을 가지고 있다.

$$\begin{aligned} A &= P \Lambda P^{-1} \\ &= P \Lambda P^T \quad \text{--- (24)} \\ \text{단, } PP^T &= E \quad (P^{-1} = P^T) \end{aligned}$$

즉, 모든 대칭행렬은 위 식 (24)와 같이 직교행렬을 이용한 고유값 대각화가 가능하다 (증명1, 증명2).

☞ 위 성질은 원소값이 실수(real number)인 경우에 항상 성립하는 성질이며 만일 원소값이 복소수(complex number)인 경우에는 유니터리(unitary) 행렬로 항상 대각화 가능하다. 그런데 사실 직교행렬을 복소수 공간에서 정의한 것이 유니터리 행렬이기 때문에 대칭행렬은 항상 대각화 가능하다고 생각하면 된다.

이렇게 따로 대칭행렬의 대각화에 대해 단락을 구분하여 적는 이유는 선형대수학에서 워낙 중요한 성질이기 때문이다. 대칭행렬(symmetric matrix)의 이러한 성질은 특이값분해(SVD), 주성분분석(PCA) 등에서 가장 기본이 되는 성질로 활용된다.

모든 정방행렬이 고유값 분해가 가능한 것은 아니지만 대칭행렬은 항상 고유값 분해가 가능하며 더구나 직교행렬로 대각화가 가능함을 기억하자.

### ※ 직교(orthogonal)와 정규직교(orthonormal), 그리고 직교행렬(orthogonal matrix)

orthogonal과 orthonormal이 서로 용어가 혼동되기 쉬운데 두 용어의 차이를 명확히 할 필요가 있다.

먼저, 벡터에 대해 얘기를 해 보면 두 벡터  $v_1, v_2$ 가 서로 수직이면(즉,  $v_1 \cdot v_2 = 0$ ) 두 벡터  $v_1, v_2$ 는 서로 orthogonal하다고 한다. 그리고  $v' = v / \|v\|$ 와 같이 어떤 벡터를 크기가 1인 단위벡터로 만드는 것을 정규화(normalization)라고 한다. orthonormal이라는 말은 orthogonal과 normal이 합쳐진 말로서 두 벡터  $v_1, v_2$ 가 모두 단위벡터(unit vector)이면서 서로 수직이면 두 벡터  $v_1, v_2$ 는 orthonormal(정규직교)하다고 한다.

- orthogonal:  $v_1 \cdot v_2 = 0$
- orthonormal:  $v_1 \cdot v_2 = 0$  &  $\|v_1\| = 1, \|v_2\| = 1$

즉, orthogonal, orthonormal은 벡터들 사이의 관계를 나타내는 말인데, 이게 행렬로 넘어가면 조금 의미가 달라진다.

먼저, 행렬에서는 직교행렬(orthogonal matrix)이라는 말은 있어도 정규직교행렬(orthonormal matrix)이라는 말은 없다. 흔히 orthonormal matrix라는 표현을 쓰는데 이는 잘못된 것이며 orthogonal matrix (직교행렬)가 올바른 용어이다.

직교행렬(orthogonal matrix)의 수학적 정의는 자신의 전치행렬(transpose)을 역행렬로 갖는 정방행렬이다.

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

$$AA^T = E \quad \text{--- (25)}$$

이와 같이 직교행렬(orthogonal matrix)은 transpose를 시키면(행렬의 열과 행 원소들을 서로 바꾸면) 자신의 역행렬이 되기 때문에 다양한 선형대수학 계산에서 매우 편리한 성질을 가진 행렬이다.

그런데, 직교행렬의 열벡터들은 서로 orthonormal(정규직교)한 성질을 가지고 있다. 즉, 직교 행렬을 구성하는 열벡터들을  $v_1, v_2, \dots, v_n$ 이라 했을 때 이들은 모두 단위벡터(unit vector)이면서 또한 서로 수직인 성질을 갖는다. 이는 식 (25)로부터 쉽게 도출될 수 있다.

$$\|v_i\| = 1, \quad i=1, \dots, n$$

$$v_i \cdot v_j = 0, \quad i \neq j \quad \text{--- (26)}$$

이러한 성질은 열벡터가 아닌 행벡터들에 대해서도 동일하게 성립한다 (즉, 행벡터들도 서로 orthonormal 하다).

즉, 직교행렬(orthogonal matrix)은 그 행렬을 구성하는 열벡터(행벡터)들이 서로 수직(orthogonal)이면서 크기가 1인 (normal한) 행렬로도 정의될 수 있다.

이상의 내용을 정리하면 다음과 같다.

- 벡터들이 orthogonal하다: 서로 수직이다
- 벡터들이 orthonormal하다: 서로 수직이면서 크기가 1인 단위벡터이다
- 행렬이 orthogonal하다:  $AA^T = E$  이다 (행렬을 구성하는 열벡터, 또는 행벡터들이 orthonormal하다)

[선형대수학 #1] 주요용어 및 기본공식

[선형대수학 #2] 역행렬과 행렬식(determinant)

[선형대수학 #3] 고유값과 고유벡터 (eigenvalue & eigenvector)

[선형대수학 #4] 특이값 분해(Singular Value Decomposition, SVD)의 활용

[선형대수학 #5] 선형연립방정식 풀이