공분산(Covariance)

어떤 두 가지 변수(x와 y)가 변하는 정도를 수로 나타낸 것

ex) x가 변할 때 y가 변하는 정도, 두 변수가 동시에 변하는 정도

두 변수가 동시에 평균에서 벗어난 정도

평균 $E(X) = \frac{\sum_{i=1}^{N} x_i}{N} = \sum x_i p_i(x_i) = \mu = \bar{x}$

분산 $V(X) = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N} = E\{(X - \mu)^2\}$

공분산 $\mathrm{Cov}(X, Y) = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{N} = E(X\text{-}\mu)\ (Y\text{-}\nu)$

공분산(Covariance)

	국어	영어
A	60	80
B	80	90

$$\text{국어의 평균} = \frac{60+80}{2} = 70, \quad \text{분산} = \frac{(60-70)^2 + (80-70)^2}{2} = 100$$

$$\text{영어의 평균} = \frac{80+90}{2} = 85, \quad \text{분산} = \frac{(80-85)^2 + (90-85)^2}{2} = 25$$

$$\text{국어·영어 공분산} = \frac{(60-70)(80-85) + (80-70)(90-85)}{2} = 50$$

공분산 행렬(Covariance Matrix)

	국어	영어
국어	100	50
영어	50	25

예제) 국영수 점수

A(90, 80, 90), B(70, 70, 50)

의 공분산 행렬을 구하시오.

고유값과 고유벡터 (eigenvalue & eigenvector)

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 벡터를 고유벡터(eigenvector)라 하고 이 상수배 값을 고유값(eigenvalue)라 한다.

즉, n x n 정방행렬(고유값, 고유벡터는 정방행렬에 대해서만 정의된다) A에 대해 Av = λv를 만족하는 0이 아닌 열벡터 **v**를 고유벡터, 상수 λ를 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 λ는 '행렬 A의 고유값', v는 '행렬 A의 λ에 대한 고유벡터'이다.

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

5. 고유값, 고유벡터의 계산

고유값과 고유벡터를 정의하는 식인 식 (1)으로 다시 돌아가서 이를 v에 대해 정리해보면 다음과 같다.

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

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

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

이 때, 식 (15)를 행렬 A의 특성방정식(characteristic equation)이라고 부르며 식 (15)를 λ에 대해 풀면 A의 고유값을 구할 수 있다. 고유벡터는 이렇게 구한 λ를 다시 식 (14)에 대입하여 계산한다.

이제 어떤 선형연산자로써의 행렬이 있을 때, 이 행렬로부터 고유 값과 고유벡터를 계산해 냄으로써, 이 행렬이 벡터에 가하는 변환을 고유벡터들이 기저를 이루는 고유공간에서의 변환으로 해석함으로써, 회전 변환은 배제하고 확대/축소 변환만으로 이해하고 응용할 수 있게 되었습니다.

확실한 이해를 위해서 구체적인 예를 한 가지만 들어 보겠습니다. 다음과 같은 행렬 A가 있습니다.

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \longrightarrow$$ 계산해봅시다.

위 행렬 A의 고유 값은 $\lambda_1 = 3$, $\lambda_2 = -1$, 고유벡터는 $v_1 = (1, 1)$, $v_2 = (1, -1)$입니다.

그러면 이 행렬 A가 벡터 $\mathbf{x} = (x_1, x_2)$ 에 변환을 가할 때, 이 변환을 고유벡터 v_1, v_2가 기저를 이루어 생성하는 고유벡터공간에서 이해를 하면, v_1 방향으로는 3배, v_2 방향으로는 -1배 한 벡터들의 합으로 변환이 가해진다는 것을 알 수 있습니다.

$$A=\begin{pmatrix}1&2\\2&1\end{pmatrix},\quad \det\left(A-\lambda E\right)=0$$

$$\begin{pmatrix}1&2\\2&1\end{pmatrix}-\begin{pmatrix}\lambda&0\\0&\lambda\end{pmatrix}=\begin{pmatrix}1-\lambda&2\\2&1-\lambda\end{pmatrix}$$

$$\det\begin{pmatrix}1-\lambda&2\\2&1-\lambda\end{pmatrix}=(1-\lambda)^2-4=0$$

$$\lambda=3,-1$$

$$\lambda=3,\quad \begin{pmatrix}1&2\\2&1\end{pmatrix}-\begin{pmatrix}3&0\\0&3\end{pmatrix}=\begin{pmatrix}-2&2\\2&-2\end{pmatrix}$$

$$\begin{pmatrix}-2&2\\2&-2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix},\quad x=y,\quad \mathbf{v}_1=\begin{pmatrix}1\\1\end{pmatrix}$$

$$\lambda=-1,\quad \begin{pmatrix}1&2\\2&1\end{pmatrix}-\begin{pmatrix}-1&0\\0&-1\end{pmatrix}=\begin{pmatrix}2&2\\2&2\end{pmatrix}$$

$$\begin{pmatrix}2&2\\2&2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix},\quad x=-y,\quad \mathbf{v}_2=\begin{pmatrix}-1\\1\end{pmatrix}$$

예제) 다음 행렬의 eigen value와 eigen vector를 구하고 해석하시오.

$$A=\begin{pmatrix}1&1\\1&1\end{pmatrix}$$

[예]

다음과 같은 행렬 A에 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 A를 식 (14)에 대입하여 특성다항식을 구해보면

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

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

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

이제 특성방정식의 해는 λ = 1, 2인데 잘 보면 λ=2는 단일근임에 비해 λ=1은 이중근임을 알 수 있다. λ에 대응하는 고유벡터의 개수는 λ가 몇중근이냐와 밀접한 관계가 있는데 단일근에 대해서는 1개, 이중근에 대해서는 최대 2개, 삼중근에 대해서는 최대 3개, ... 의 고유벡터가 존재한다.

먼저 λ = 2를 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, λ = 2에 대응하는 고유벡터는 v = [1, 1, 0]^T 이다.

마찬가지로, λ = 1에 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 λ = 1에 대응하는 고유벡터를 [2, 0, 1]과 [0, 1, 0] 으로 잡을 수 있다.

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

식(1)의 양변에 상수 c를 곱해보면,

$$A(c\mathbf{v})=\lambda(c\mathbf{v}) \quad \text{---} \ (21)$$

가 되므로 **v**가 λ에 대한 고유벡터이면 0이 아닌 임의의 상수 c에 대해 c**v**도 또한 λ에 대한 고유벡터임을 알 수 있다.

또한, v1, v2가 모두 고유값 λ에 대응하는 고유벡터라고 하면 임의의 상수 c1, c2에 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 c1v1 + c2v2 또한 λ에 대한 고유벡터가 됨을 알 수 있다.

Determinant

3x3 이상의 행렬

이 때에는 다소 생소한 방법으로 행렬식을 구해야 합니다.[2] 3x3 이상의 행렬에 대한 행렬식을 구하는 방법을 익힐 때에는 새로운 개념의 행렬이 등장합니다. 바로 소행렬식(minor)과 여인자(cofactor; 여인수)라는 것입니다.

먼저 소행렬식을 설명드리겠습니다. 소행렬식은 보통 **(i,j)th minor** 이렇게 표현하는데, i번째 행과 j번째 열을 제거한 새로운 행렬의 행렬식을 말합니다. 예를 들어 3x3 행렬 B를 봅시다.

$$B=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

이 B 행렬의 (2,3)th minor라고 하면 2번째 행과 3번째 열을 제거한

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix}$$ 행렬의 행렬식을 말하는 것입니다.

$$(i,j)th\ minor = M_{ij} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

이렇게 만들 수 있는 소행렬식은 (1,1)부터 (3,3)까지 총 9가지가 되겠습니다. (4x4행렬이었다면 16개가 되었겠네요.)

이제 여인자(cofactor)라는 것을 알아봅시다. 이것도 역시 **(i,j)th Cofactor**라는 식으로 표현하는데요, (i,j)th minor에 $(-1)^{i+j}$를 곱한 값입니다.

$$(i,j)th\ cofactor = \mathrm{C}_{ij} = (-1)^{i+j}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

이제 3x3이상 행렬의 행렬식은 아무 행이나 아무 열 하나를 잡습니다. 그 행이나 그 열의 원소와 여인자를 곱한 것의 합을 냅니다. 예시로 위의 3x3 행렬로 보여드리겠습니다.

$$\mathrm{B} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

저는 2번째 열로 행렬식을 구해보겠습니다.

$$|\mathrm{B}| = det(\mathrm{B}) = a_{12}\mathrm{C}_{12} + a_{22}\mathrm{C}_{22} + a_{32}\mathrm{C}_{32}$$

$$|\mathrm{B}| = det(\mathrm{B}) = a_{12}(-1)^{1+2}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{22}(-1)^{2+2}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + a_{32}(-1)^{2+3}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}$$

$$|\mathrm{B}| = det(\mathrm{B}) = -a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{22}(a_{11}a_{33} - a_{13}a_{31}) - a_{32}(a_{11}a_{23} - a_{13}a_{21})$$

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \quad aei + bfg + cdh$$

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \quad ceg + bdi + afh$$

$$\det\left|B\right| = (aei + bfg + cdh) - (ceg + bdi + afh)$$