chap5. eigenvalue & eigenvector

$A\vec{x} = \lambda\vec{x}$ 이 성립하는 경우, $\vec{x}$를 eigenvector, λ를 eigenvalue라고 한다.
(A_{n*n} 이고 λ는 scalar 값)

그렇다면 고유값과 고유벡터는 어떻게 구할까?

$(A-\lambda I)\vec{x} = 0$ 이므로 for non-zero $\vec{x}$, $det(A-\lambda I) = 0$ 이여야 한다. det를 통해 우리는 고유값(λ)을 구할 수 있고 이에 해당하는 고유벡터를 구할 수 있다. $A - \lambda I$의 null space를 이루는 벡터들이다.
고유값은 n개가 존재하며 가가 고유값마다 해당하는 고유벡터가 하나씩 존재한다. 물론 고유값에 중복이 존재할 수 있다. 이런 경우는 n개가 되지 않는다.

triangular (or diagonal) Matrix

이 행렬의 고유값은 대각성분들이다. A가 triangular, diagonal 행렬일 때,

$$\begin{bmatrix} a_{11}-\lambda & a_{12} & ... & a_{1n} \\ 0 & a_{22}-\lambda & ... & a_{2n} \\ ... & ... & ... & ... \\ 0 & 0 & ... & a_{nn}-\lambda \end{bmatrix}$$

위 행렬에서 $det(A-\lambda I) = (a_{11}-\lambda)(a_{22}-\lambda)...(a_{nn}-\lambda) = 0$ 이므로 고유값들은 대각성분들이 되는 것이다. 따라서 고유값들의 곱 = $det(A)$

trace of A

trace of A는 행렬 A 대각성분의 합을 의미한다. 행렬 A 대각성분의 합은 그 행렬의 고유값의 합과 같다.

$$\sum_{i=1}^{n} a_{ij} = \sum_{i=1}^{n} \lambda_i$$

diagonalization of matrix

이전에 배운 분해는
(1) $A = LU = LDU$

(2) $A = QR$

이번에는 고유값을 이용하여 분해를 해보자
(3) $A = S\Lambda S^{-1}$

이제 자세히 알아보자.

(e_i : eigenvector), $A\vec{e}_i = \lambda_i\vec{e}_i$ 이므로

$$\begin{bmatrix} A\vec{e}_1 & A\vec{e}_2 & .... & A\vec{e}_n \end{bmatrix} = \begin{bmatrix} \lambda\vec{e}_1 & \lambda\vec{e}_2 & .... & \lambda\vec{e}_n \end{bmatrix}$$

$$= A\begin{bmatrix} \vec{e}_1 & \vec{e}_2 & .... & \vec{e}_n \end{bmatrix} = \begin{bmatrix} \vec{e}_1 & \vec{e}_2 & .... & \vec{e}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{bmatrix}$$

위와 같은 모습은 $AS = S\Lambda$ 이다. 따라서 $A = S\Lambda S^{-1}$

remark

(1) if $\lambda_1, \lambda_2, ..., \lambda_n$ are different, then $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$ are linearly independent

앗 그렇다면 새로운 basis로 사용 가능하겠구나! 라는 생각을 할 수 있다.
(proof)
assume that $\vec{e}_2 = c\vec{e}_1$, 양변에 각각 A 와 λ_2를 곱하면

- $A\vec{e}_2 = cA\vec{e}_1$ -> $\lambda_2\vec{e}_2 = c\lambda_1\vec{e}_1$
- $\lambda_2\vec{e}_2 = c\lambda_2\vec{e}_1$

위 두 식에 의해서 $c\lambda_1\vec{e}_1 = c\lambda_2\vec{e}_1$ 이므로 $c(\lambda_2 - \lambda_1)\vec{e}_1 = 0$ 이다. 그런데 고유값은 서로 다르고 $\vec{e}$는 not zero vector이므로 $c = 0$ 이여야 한다. 이런식으로 하면 모든 고유벅벡터들은 linearly independent이다.

(2) 위에서 봤던 S는 unique하지 않다. eigenvector들의 모양이 달라질 수 있기 때문이다. 하지만 constant가 곱해진 eigenvector가 달라도 최종적으로 행렬 A가 만들어지는 것은 똑같다.

(3) the order of eigenvalues is same with that of eigenvectors

(4) not all matrix have n linearly independent eigenvector, 고유값이 중복될 수 있기 때문이다. 이럴 경우 $S\Lambda S^{-1}$ 도 불가능하다.

least square $A\vec{x} = 0$

$A\vec{x} = \vec{b}$에서 $\vec{b}$가 0이 아니면 이전의 배운 least sqaure 방법으로 $\vec{x}$를 구할 수 있다. 하지만 b가 0인 homogeneous equation의 경우는 어떻게 해야 할까? (LS로 구한 $\vec{x} = (A^TA)^{-1}A^T\vec{b}$에서 $\vec{b} = \vec{0}$이면 $\vec{x} = 0$밖에 없음)

$min \left\| A\vec{x} - 0 \right\|^2$ 에서 $\vec{x}$가 워낙 다양하게 나오니까 $\left\| \vec{x} \right\| = 1$로 제한

$$\left\| A\vec{x} \right\|^2 = (A\vec{x})^T(A\vec{x}) = \vec{x}^TA^TA\vec{X}$$

위에서 A^TA는 sqaure matrix이다. 이를 B라고 한다면 $B\vec{e} = \lambda\vec{e}$을 이용하면

$$\left\| A\vec{x} \right\|^2 = \vec{x}^TB\vec{x} = \vec{x}^T\lambda\vec{x} = \lambda\left\| \vec{x} \right\|^2 = \lambda$$

즉, $\left\| A\vec{x} \right\|^2$의 가장 작은 값은 A^TA의 가장 작은 고유값이고 결국 이 때의 해 $\vec{x}$는 그에 해당하는 고유벡터인 것이다.

- $min \left\| A\vec{x} - 0 \right\|^2$ subject to $\vec{x}$ $\left\| \vec{x} \right\| = 1$

$J(\vec{x}, \lambda) = \left\| A\vec{x} \right\|^2 + \lambda(1 - \left\| \vec{x} \right\|^2)$ 위의 조건을 하나의 식으로 표현할 수 있다. 여기서 λ를 Lagrange multiplier 라고 한다. 위의 J에 대해 각각 $\vec{x},\ \lambda$로 편미분하여 $min\ J$ 하게 해주는 $\vec{x},\ \lambda$ 구할 수 있다.

Markov matrix

어떠한 transition을 나타내는 행렬이라고 할 수 있다. 예를 들어 이해보자. 매년 A지역과 B지역의 인구변동이 있다고 하자.
(1) A지역에서 B지역으로 이사가는 비율이 0.1, A지역에 머무르는 비율이 0.9

(2) B지역에서 A지역으로 이사가는 비율이 0.2, B지역에 머무르는 비율이 0.8
이러한 상황이 있다고 하고 이를 행렬로 나타내보자.

$$\begin{bmatrix} y_{k+1} \\ x_{k+1} \end{bmatrix} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} y_k \\ x_k \end{bmatrix}$$

위 식은 $U_{k+1} = AU_k$의 형태를 가지고 결국

$$U_{k+1} = A^k U_0 = S\Lambda^k \Lambda^{-1} U_0 = \begin{bmatrix} \vec{e}_1 & \vec{e}_2 \end{bmatrix} \begin{bmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

k가 무한대로 가면 U_k가 converge하는 경우 이 값을 stable state라고 한다.

이런 Markov matrix의 특징은

(1) column 마다 각 원소들은 0 과 1 사이의 값을 가지고 원소들의 합은 1 이다.

(2) 하나의 eigenvalue 는 1이고 나머지는 절댓값이 1보다 작다.

difference equation & A^k

피보나치 수열을 예시로 드셨다.

미분방정식

여기 내용은 skip하겠다

complex matrix

$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{bmatrix}$ 일 때, $x_i = a_k + jb_k$ when $j^2 = -1$ 이런 벡터를 complex vector 라고 한다. 이전까지는 실수로 이루어진 결우만 생각했다면 이제 복소수까지로 개념을 확장해보는 것이다.

$$\|\vec{x}\|^2 = {x_1}^* x_1 + {x_2}^* x_2 + ... + {x_n}^* x_n = (\vec{x}^T)^* \vec{x}$$

여기서 * 표시는 켤레복소수를 의미한다. 그리고 위에서 transpose와 conjugate를 동시에 하는 hermite라고 하며 이제부터 H라고 표시한다.

- inner product of complex vector
  real vector : $\vec{x}^T \vec{y} = \vec{y}^T \vec{x}$
  complex vector : $\vec{x}^H \vec{y} \neq \vec{y}^H \vec{x}$

(1) if $\vec{x}^H y =$ real , $\vec{x}^H \vec{y} = \vec{y}^H \vec{x}$ 따라서 vector가 orthogonal하면 $\vec{x}^H \vec{y} = \vec{y}^H \vec{x} = 0$

(2) $\|\vec{x}\|^2 = \vec{x}^H \vec{x}$

(3) $(AB)^H = B^H A^H$

Hermitian matrix

이전에

- real matrix의 경우에서 $A^T = A$: symmetric matrix
- complex matrix의 경우에서 $A^H = A$: hermitian matrix 라고 한다. $a_{ij} = a^*_{ji}(i \neq j)$, 대각성분은 real 이다.

여기서 hermitian 이 좀 더 큰 범위인 것이다.

- properties of H or S matrix : A

(1) $\vec{x}^H A\vec{x}$ (quadratic form) is real :
$(\vec{x}^H A\vec{x})^H = \vec{x}^H A^H \vec{x} = \vec{x}^H A\vec{x}$ 이기 때문이다. (H를 취했는데 같은 값이 나오면 real임을 의미)

(2) A의 모든 고유값은 real 이다 :
$A\vec{e} = \lambda\vec{e}$ 에서 $\vec{e}^H A\vec{e} = \vec{e}^H \lambda\vec{e} = \lambda \|\vec{e}\|^2$ 이다. 따라서 $\lambda = \frac{\vec{e}^H A\vec{e}}{\|\vec{x}\|^2}$ (1)에서 분자가 real이고 분모도 real이므로 λ는 real 이다.

(3) 고유벡터들이 orthorgonal (orthornormal) 하다 :
이전까지는 고유벡터가 independent 했는데 애는 orthorgonal하다!
$A\vec{x}_1 = \lambda_1\vec{x}_1,\ A\vec{x}_2 = \lambda_2\vec{x}_2, (\lambda_1 \neq \lambda_2)$

- $(A\vec{x}_1)^H\vec{x}_2 = \vec{x}_1^H A^H \vec{x}_2 = \vec{x}_1^H A\vec{x}_2 = \lambda_2\vec{x}_1^H\vec{x}_2$
- $(A\vec{x}_1)^H\vec{x}_2 = (\lambda_1\vec{x}_1)^H\vec{x}_2 = \lambda_1^*\vec{x}_1^H\vec{x}_2 = \lambda_1\vec{x}_1^H\vec{x}_2$

따라서 $\lambda_2\vec{x}_1^H\vec{x}_2 = \lambda_1\vec{x}_1^H\vec{x}_2$ 이고 $(\lambda_2 - \lambda_1)\vec{x}_1^H\vec{x}_2 = 0$ 이다. 두 고유값은 다르기 때문에 $\vec{x}_1^H\vec{x}_2 = 0$ 이고 고유벡터들은 orthorgonal하다.

spactral them

A: symmetric matrix (고유벡터는 길이 1로 normalize)

$A = S\Lambda S^{-1} = Q\Lambda Q^{-1}$ (A의 고유벡터들이 orthonormal 하기 때문에 이전에 배웠던 orthonormal matrix Q로 표현가능하다)
$= Q\Lambda Q^T$ ($Q^TQ = I$ 이므로)

$$= \begin{bmatrix} \vec{e}_1 & \vec{e}_2 & .... & \vec{e}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{bmatrix} \begin{bmatrix} \vec{e}_1^T \\ \vec{e}_2^T \\ ... \\ \vec{e}_n^T \end{bmatrix}$$

$$A = \lambda_1\vec{e}_1\vec{e}_1^T + \lambda_2\vec{e}_2\vec{e}_2^T + ... + \lambda_n\vec{e}_n\vec{e}_n^T$$

A라는 행렬(space)은 고유벡터들의 projection 성분으로 이루어져 있고 가중치는 그에 해당하는 고유값이다. (projection 공부했던 부분에서 $P = \frac{\vec{e}_i\vec{e}_i^T}{\vec{e}_i\vec{e}_i}$ 인데 orthonormal이라 분모가 1)

Unitary Matrix : U (Q의 확장판)

Q랑 똑같은데 complex matrix이다.

$$U^HU = UU^H = I,\ U^H = U^{-1}$$

(1) angle, length preserved

length : $\|U\vec{x}\|^2 = (U\vec{x})^T(U\vec{x}) = \|\vec{x}\|^2$
angle : $(U\vec{x})^H(U\vec{y}) = \vec{x}^H\vec{y}$

(2) 고유값들의 길이가 1이다

$|\lambda| = 1$,
$U\vec{e} = \lambda\vec{e}$ 을 이용하여 $\|U\vec{x}\| = \|\lambda\vec{x}\| = |\lambda|\,\|\vec{x}\|$ 그런데 위에서 $\|U\vec{x}\| = \|\vec{x}\|$ 임을 알았다. 따라서 $|\lambda| = 1$

(3) 고유벡터들이 orthonormal 하다

$U\vec{e}_1 = \lambda_1\vec{e}_1,\ U\vec{e}_2 = \lambda_2\vec{e}_2, (\lambda_1 \neq \lambda_2)$ 이라고 하면

$$\vec{x}_1^H\vec{x}_2 = (U\vec{x}_1)^H(U\vec{x}_2) = (\lambda_1\vec{x}_1)^H(\lambda_2\vec{x}_2) = \lambda_1^*\lambda_2\vec{x}_1^H\vec{x}_2$$

$$(1 - \lambda_1^*\lambda_2)\vec{x}_1^H\vec{x}_2 = 0$$

λ의 길이가 1인 상태에서 $\lambda_1^*\lambda_2 = 1$이 되기 위해서는 $i = j$ 인 경우에만 1 이기 때문에

$$\vec{x}_1^H\vec{x}_2 = 0$$