MATRIX DECOMPOSITION

Homepage: https://sites.google.com/site/doc4code/

Email: sj6219@hotmail.com

2012/10/08

이 문서에서는 행렬에 관련된 수학 법칙에 대해서 공부해 보겠다.

목차

1 QR DECOMPOSITION

QR decomposition 이란 matrix A 를 unitary matrix Q와 upper triangular matrix R의 곱으로 분해하는 것은 말한다.

$$A \in \mathbb{C}^{m \times n}, Q \in \mathbb{C}^{m \times m}, R \in \mathbb{C}^{m \times n}$$

$$A = QR$$

여기서, unitary matrix 란

$$Q \in \mathbb{C}^{m \times m}$$

$$Q^*Q = QQ^* = I$$

을 만족하는 Q를 말한다.

만약, $A \in \mathbb{R}^{m \times n}$이라면 $Q \in \mathbb{R}^{m \times m}, R \in \mathbb{R}^{m \times n}$ 으로 분해할 수 있다.

QR decomposition 에는 Gram-Schmidt, Householder Reflection, Givens rotation 방식이 주로 사용된다.

여기서는 비교적 간단한 Householder Reflection 방식에 대해서 알아보자

HOUSEHOLDER REFLECTION

어떤 n 차원 벡터 $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ 에 대하여,

$$\alpha \in \mathbb{C}$$

$$H\mathbf{x} = \alpha\, \mathbf{e_1}$$

$$\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

을 만족시키는 n×n unitary matrix H 를 vector $\mathbf{x}$ 에 대한 Householder reflection matrix 라고 부른다.

이해를 돕기 위해서 $\mathbf{x} \in \mathbb{R}^n$인 경우부터 알아보겠다.

$$\mathbf{u} \in \mathbb{R}^n, \mathbf{v} \in \mathbb{R}^n$$

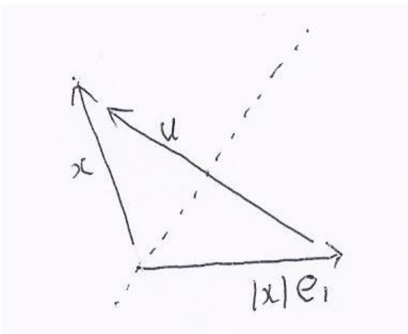
$$\alpha = |\mathbf{x}| \; or \; -|\mathbf{x}|$$

$$\mathbf{u} = \mathbf{x} - \alpha\, \mathbf{e_1}$$

$$\mathbf{v} = \frac{\mathbf{u}}{|\mathbf{u}|}$$

$$H = I - 2\mathbf{v}\mathbf{v}^T$$

를 이용해 H 를 구할 수 있다.



기하학적으로 보면 H 는 $\mathbf{x}$ 를 $\mathbf{e_1}$ 방향으로 변환하는 평면 대칭 이동이다.

이번엔 좀 더 복잡한 $\mathbf{x} \in \mathbb{C}^n$ 인 경우에 대해 구해보자.

α 는 크기가 $|\mathbf{x}|$ 인 임의의 복소수(complex)로 설정하고,

$$\mathbf{u} \in \mathbb{C}^n, \mathbf{v} \in \mathbb{C}^n$$

$$|\alpha| = |\mathbf{x}|$$

$$\mathbf{u} = \mathbf{x} - \alpha\,\mathbf{e_1}$$

$$\mathbf{v} = \frac{\mathbf{u}}{|\mathbf{u}|}$$

$$\mathrm{H} = \mathrm{I} - \left(1 + \frac{\mathbf{x}^*\mathbf{v}}{\mathbf{v}^*\mathbf{x}}\right)\mathbf{v}\mathbf{v}^*$$

를 이용해 H 를 구할 수 있다.

$$\mathrm{x}_1 = \mathbf{e_1}^{\mathrm{T}}\,\mathbf{x}$$

$$\mathbf{u}^*\mathbf{x} = \mathbf{x}^*\mathbf{x} - \bar{\alpha}\,\mathrm{x}_1 = |\mathbf{x}|^2 - \bar{\alpha}\,\mathrm{x}_1$$

$$\mathbf{x}^*\mathbf{u} = \mathbf{x}^*\mathbf{x} - \alpha\,\overline{\mathrm{x}_1} = |\mathbf{x}|^2 - \alpha\,\overline{\mathrm{x}_1}$$

$$\mathbf{u}^*\mathbf{u} = \mathbf{x}^*\mathbf{x} - \bar{\alpha}\,\mathrm{x}_1 - \alpha\,\overline{\mathrm{x}_1} + |\alpha|^2 = 2|\mathbf{x}|^2 - \bar{\alpha}\,\mathrm{x}_1 - \alpha\,\overline{\mathrm{x}_1}$$

이므로

$$\mathrm{H}\mathbf{x} = \left(\mathrm{I} - \left(1 + \frac{\mathbf{x}^*\mathbf{v}}{\mathbf{v}^*\mathbf{x}}\right)\mathbf{v}\mathbf{v}^*\right)\mathbf{x}$$

$$= \mathbf{x} - \left(1 + \frac{\mathbf{x}^*\mathbf{u}}{\mathbf{u}^*\mathbf{x}}\right)\frac{\mathbf{u}^*\mathbf{x}}{\mathbf{u}^*\mathbf{u}}\,\mathbf{u}$$

$$= \mathbf{x} - \left(\frac{\mathbf{u}^*\mathbf{x} + \mathbf{x}^*\mathbf{u}}{\mathbf{u}^*\mathbf{u}}\right)\mathbf{u}$$

$$= \mathbf{x} - \mathbf{u} = \alpha\,\mathbf{e_1}$$

또,

$$\mathrm{H}^*\mathrm{H} = \left(\mathrm{I} - \left(1 + \frac{\mathbf{v}^*\mathbf{x}}{\mathbf{x}^*\mathbf{v}}\right)\mathbf{v}\mathbf{v}^*\right)\left(\mathrm{I} - \left(1 + \frac{\mathbf{x}^*\mathbf{v}}{\mathbf{v}^*\mathbf{x}}\right)\mathbf{v}\mathbf{v}^*\right) = \mathrm{I}$$

이므로 H 는 unitary matrix 이다.

지금까지 Householder reflection matrix 를 구하는 방법을 배웠다.

이제, 본격적으로 m×n matrix A 를 QR decomposition 을 해보자.

m 이 1 인 경우에는 다음과 같이 unitary matrix 1 과 upper triangular matrix A로 분해된다.

$$\mathrm{A} = 1\,\mathrm{A}$$

m 이 2 이상인 경우에

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

로 표현될 때,

$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$에 대한 Householder reflection matrix 를 H이라고 하자.

$$H \in \mathbb{C}^{m \times m}$$

$$HA = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

가 된다.

그런데, 이 행렬을 다음과 같이 분할(partition)할 수 있다.

$$\alpha \in \mathbb{C}, \boldsymbol{\beta} \in \mathbb{C}^{n-1}, A_1 \in \mathbb{C}^{(m-1)\times(n-1)}$$

$$HA = \begin{bmatrix} \alpha & \boldsymbol{\beta}^* \\ 0 & A_1 \end{bmatrix}$$

재귀적으로 A_1를 QR decomposition 을 해서,

$$Q_1 \in \mathbb{C}^{(m-1)\times(m-1)}, R_1 \in \mathbb{C}^{(m-1)\times(n-1)}$$

$$A_1 = Q_1 R_1$$

로 분해되었다면,

$$\begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix}^* HA = \begin{bmatrix} \alpha & \boldsymbol{\beta}^* \\ 0 & {Q_1}^* A_1 \end{bmatrix} = \begin{bmatrix} \alpha & \boldsymbol{\beta}^* \\ 0 & R_1 \end{bmatrix}$$

$$A = H^* \begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix} \begin{bmatrix} \alpha & \boldsymbol{\beta}^* \\ 0 & R_1 \end{bmatrix}$$

이 되므로, A는 unitary matrix $Q = H^* \begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix}$ 와 upper triangular matrix $R = \begin{bmatrix} \alpha & \boldsymbol{\beta}^* \\ 0 & R_1 \end{bmatrix}$ 로 분해된다.

2 SCHUR DECOMPOSITION

Schur decomposition 은 square matrix A를 아래와 같이 unitary matrix Q 와 upper triangular matrix T 로 분해하는 것이다.

$$\mathrm{A} \in \mathbb{C}^{n\times n}, \mathrm{Q} \in \mathbb{C}^{n\times n}, \mathrm{T} \in \mathbb{C}^{n\times n}$$

$$\mathrm{A} = \mathrm{QTQ}^*$$

그 방법에 대해 알아보자.

n 이 1 일 때는 A 는

$$\mathrm{A} = 1\,\mathrm{A}\,1^*$$

로 간단히 분해된다.

n 이 2 이상일 때에는 A 의 한 개 이상의 eigenvalue λ_1와 그 eigenvector $\mathbf{x_1}$ 가 존재한다.

eigenvector $\mathbf{x_1}$ 에 수직(orthogonal)인 벡터들을 구하기 위해, $\mathrm{n} \times 1$ matrix 로 취급하여 아래와 같이 QR decomposition 해서 unitary matrix U 를 구한다.

$$\mathbf{x_1} \in \mathbb{C}^{n\times 1}, \mathrm{U} \in \mathbb{C}^{n\times n}, \alpha \in \mathbb{C}$$

$$\mathbf{x_1} = \mathrm{U}\begin{bmatrix}\alpha \\ 0 \\ \vdots \\ 0\end{bmatrix}$$

이 때,

$$|\alpha| = |\mathbf{x_1}|$$

이어야 한다.

또한, U은 아래처럼 분할되므로

$$\mathrm{U} = \begin{bmatrix}\frac{\mathbf{x_1}}{\alpha} & * & \cdots & *\end{bmatrix}$$

$$\mathrm{U}^*\mathrm{AU} = \mathrm{U}^*\begin{bmatrix}\frac{\mathrm{A}\mathbf{x_1}}{\alpha} & * & \cdots & *\end{bmatrix}$$

$$= \mathrm{U}^*\begin{bmatrix}\frac{\lambda_1\mathbf{x_1}}{\alpha} & * & \cdots & *\end{bmatrix}$$

U의 첫번째 column vector $\frac{\mathbf{x_1}}{\alpha}$는 다른 column vector 와 수직(orthogonal)이므로

$$\mathrm{U}^*\mathrm{AU} = \begin{bmatrix}\lambda_1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & *\end{bmatrix}$$

U^*AU 를 다음과 같이 분할하면

$$B \in \mathbb{C}^{1\times(n-1)}, A_1 \in \mathbb{C}^{(n-1)\times(n-1)}$$

$$U^*AU = \begin{bmatrix} \lambda_1 & B \\ 0 & A_1 \end{bmatrix}$$

A_1를 재귀적으로 Schur decomposition 해서

$$Q_1 \in \mathbb{C}^{(m-1)\times(m-1)}, T_1 \in \mathbb{C}^{(m-1)\times(n-1)}$$

$$A_1 = Q_1 T_1 {Q_1}^*$$

로 분해되었다면,

$$\begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix}^* U^*AU \begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & BQ_1 \\ 0 & {Q_1}^* A_1 Q_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & BQ_1 \\ 0 & T_1 \end{bmatrix}$$

$$A = U \begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix} \begin{bmatrix} \lambda_1 & BQ_1 \\ 0 & T_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix}^* U^*$$

이 되므로, A 는 unitary matrix $Q = U \begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix}$ 와 upper triangular matrix $T = \begin{bmatrix} \lambda_1 & BQ_1 \\ 0 & T_1 \end{bmatrix}$ 로 분해된다.

여기서 matrix A 의 eigenvalu 가 $\lambda_1, \lambda_2, \cdots, \lambda_n$이면, matrix A_1의 eigenvalue 는 $\lambda_2, \cdots, \lambda_n$ 이라는 것을 증명해보자.

matrix A 의 eigenvalue 가 $\lambda_1, \lambda_2, \cdots, \lambda_n$라면 x 에 대한 n 차 방정식

$$\det(A - x\, I) = 0$$

와

$$(\lambda_1 - x)(\lambda_2 - x) \cdots (\lambda_n - x) = 0$$

는 같은 방정식이다.

$$\det(A - x\, I) = \det(U^*(A - x\, I)U) = \det(U^*AU - xU^*IU)$$

$$= \det\left(\begin{bmatrix} \lambda_1 & B \\ 0 & A_1 \end{bmatrix} - x\, I\right)$$

$$= \det\left(\begin{bmatrix} \lambda_1 - x & B \\ 0 & A_1 - xI \end{bmatrix}\right)$$

$$= (\lambda_1 - x)\det(A_1 - x\, I) = 0$$

위 식의 해가 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 이므로 matrix A_1의 eigenvalue 는 $\lambda_2, \cdots, \lambda_n$는 이어야 한다.

지금까지 증명한 것을 정리하면, matrix A 의 eigenvalue 가 $\lambda_1, \lambda_2, \cdots, \lambda_n$라면, A 를 다음과 같이 원하는 eigenvalue 순서대로 분해할 수 있다.

$$A = Q \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} Q^*$$

그리고, matrix A 를 Schur decomposition 했을 때, upper triangular matrix 의 대각선 원소들은 A의 eigenvalue 여야 한다.

3 NORMAL MATRIX

$AA^* = A^*A$을 만족하는 square matrix 를 normal matrix 라고 정의한다.

square matrix A를 다음과 같이 invertible matrix S와 diagonal matrix D로 분해하는 것을 eigen decomposition 이라고 부른다.

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

모든 square matrix 에 대하여 eigen decomposition 이 가능한 것은 아니다. normal matrix 의 경우에는, 항상 eigen decomposition 이 가능하다.

A 가 normal matrix 라면, 아래와 같이 Schur decomposition 할 수 있다

$$A = QTQ^*$$

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & t_{nn} \end{bmatrix}$$

$$AA^* = Q \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & t_{nn} \end{bmatrix} Q^*Q \begin{bmatrix} \overline{t_{11}} & 0 & \cdots & 0 \\ \overline{t_{12}} & \overline{t_{22}} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \overline{t_{1n}} & \overline{t_{2n}} & \cdots & \overline{t_{nn}} \end{bmatrix} Q^*$$

$$= Q \begin{bmatrix} |t_{11}|^2 + |t_{12}|^2 \cdots |t_{1n}|^2 & * & \cdots & * \\ * & |t_{22}|^2 \cdots |t_{2n}|^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ * & \cdots & * & |t_{nn}|^2 \end{bmatrix} Q^*$$

이고,

$$A^*A = Q\begin{bmatrix} \overline{t_{11}} & 0 & \cdots & 0 \\ \overline{t_{12}} & \overline{t_{22}} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \overline{t_{1n}} & \overline{t_{2n}} & \cdots & \overline{t_{nn}} \end{bmatrix} Q^*Q \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & t_{nn} \end{bmatrix} Q^*$$

$$= Q\begin{bmatrix} |t_{11}|^2 & * & \cdots & * \\ * & |t_{12}|^2 + |t_{22}|^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ * & \cdots & * & |t_{1n}|^2 + |t_{2n}|^2 \cdots |t_{nn}|^2 \end{bmatrix} Q^*$$

두 식이 같기 위해서는 T 는 diagonal matrix 여야 한다.

그러므로, normal matrix A 는 다음과 같이 unitary matrix Q 와 diagonal matrix D 로 eigen decomposition 된다.

$$A = QDQ^*$$

이 때, D 의 대각선 원소는 A 의 eigenvalue 들이다.
역으로, 모든 unitary matrix Q 와 diagonal matrix D 에 대해 QDQ^* 는 normal matrix 이다.

REAL NORMAL MATRIX

real matrix A가 normal matrix 이면, A 는 다음과 같이 분해될 수 있다.

$$A = Q\begin{bmatrix} \lambda_1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \lambda_p & \ddots & & \vdots \\ \vdots & & \ddots & A_1 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & A_q \end{bmatrix} Q^*$$

Q는 real orthogonal matrix 이고,

$$\lambda_i \in \mathbb{R}, A_i \in \mathbb{R}^{2\times 2}$$

$$A_i = \begin{bmatrix} \alpha_i & -\beta_i \\ \beta_i & \alpha_i \end{bmatrix}$$

이다.

UNITARY MATRIX

$A A^* = A^*A = I$ 을 만족하는 square matrix A 를 unitary matrix 라고 정의한다.

unitary matrix 는 normal matrix 이고, 그 eigenvalue 가 $\lambda_1, \lambda_2, \cdots, \lambda_n$라면

$$|\lambda_1| = |\lambda_2| = \cdots |\lambda_n| = 1$$

이 성립한다.

역으로, eigenvalue 의 절대값이 1 인 normal matrix 는 unitary matrix 이다.

REAL ORTHOGONAL MATRIX

real matrix A가 unitary matrix 이면 A 를 real orthogonal matrix 라고 정의한다.

A가 real orthogonal matrix 이면 다음과 같이 분해될 수 있다.

$$A = Q \begin{bmatrix} \lambda_1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \lambda_p & \ddots & & \vdots \\ \vdots & & \ddots & A_1 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & A_q \end{bmatrix} Q^*$$

Q는 real orthogonal matrix 이고,

$$\lambda_i = \pm 1$$

$$A_i = \begin{bmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix}$$

이다.

HERMITIAN MATRIX

$A^* = A$ 를 만족하는 square matrix A를 Hermitian matrix 로 정의한다.

Hermitian matrix A는 다음과 같이 분해가 가능하다.

$$A = Q \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} Q^*$$

Q는 unitary matrix 이고, $\lambda_1, \lambda_2, \cdots, \lambda_n$ 는 실수이다.

모든 Hermitian matrix A와 square matrix S에 대하여 S^*AS 도 Hermitian matrix 이다.

SYMMETRIC MATRIX

real matrix A가 Hermitian matrix 이면 A 를 symmetric matrix 라고 정의한다.

Hermitian matrix A는 다음과 같이 분해가 가능하다.

$$A = Q \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} Q^*$$

Q는 real orthogonal matrix 이고, $\lambda_1, \lambda_2, \cdots, \lambda_n$ 는 실수이다.

POSITIVE DEFINITE MATRIX

모든 0 이 아닌 n 차원 벡터 $\mathbf{x}$ 에 대하여 $\mathbf{x}^* A\, \mathbf{x} > 0$ 을 만족하는 n×n matrix A를 positive definite matrix 라고 정의한다.

$$\mathbf{x}^* A\, \mathbf{x} = \mathbf{x}^* \left(\frac{A + A^*}{2} + i \frac{A - A^*}{2i} \right) \mathbf{x} = \mathbf{x}^* \left(\frac{A + A^*}{2} \right) \mathbf{x} + i \mathbf{x}^* \left(\frac{A - A^*}{2i} \right) \mathbf{x}$$

에서, $\mathbf{x}^*\mathrm{A}\,\mathbf{x}$ 가 실수이기 위해서는 Hermitian matrix 인 $\left(\frac{\mathrm{A}-\mathrm{A}^*}{2i}\right)$ 가 0 이어야 한다. 그러므로, positive definite matrix 는 Hermitian matrix 이다. 또한, positive definite matrix 의 eigenvalue 는 모두 양의 실수이다.

A가 positive definite matrix 라면,

$$\mathrm{A} = \mathrm{B}^*\mathrm{B}$$

을 만족시키는 invertible matrix B가 존재한다.

A는 unitary matrix Q와 diagonal matrix D를 사용하여 $\mathrm{A} = \mathrm{QDQ}^*$ 로 분해할 수 있으므로, $\mathrm{B} = \mathrm{Q}\,\mathrm{D}^{\frac{1}{2}}\,\mathrm{Q}^*$인 B가 존재한다.

모든 0 이 아닌 n 차원 vector $\mathbf{x}$에 대해 $\mathrm{B}\mathbf{x} \neq 0$ 를 만족하는 m×n matrix B를 **nonsingular matrix** 라고 정의한다.

모든 nonsingular matrix B에 대하여 $\mathrm{B}^*\mathrm{B}$ 는 positive definite matrix 이다.

0 이 아닌 vector $\mathbf{x}$에 대해

$$\mathbf{y} = \mathrm{B}\mathbf{x} \neq 0$$

라고 하면,

$$\mathbf{x}^*\,\mathrm{B}^*\mathrm{B}\,\mathbf{x} = \mathbf{y}^*\mathbf{y} > 0$$

그러므로 $\mathrm{B}^*\mathrm{B}$ 는 positive definite matrix 이다.

POSITIVE SEMIDEFINITE MATRIX

모든 0 이 아닌 n 차원 벡터 $\mathbf{x}$ 에 대하여, $\mathbf{x}^*\mathrm{A}\,\mathbf{x} \geq 0$ 을 만족하는 n×n matrix A를 positive semidefinite matrix 라고 정의한다.

A가 positive semidefinite matrix 라면,

$$\mathrm{A} = \mathrm{B}^*\mathrm{B}$$

을 만족시키는 square matrix B 가 존재한다.

또한, 모든 m×n matrix B 에 대하여 B^*B는 positive semidefinite matrix 이다.

같은 방법으로 negative definite, negative semidefinite matrix 를 정의한다. 그리고, 여기에 속하지 않는 Hermitian matrix 를 indefinite matrix 라고 정의한다. Indefinite matrix 는 음수, 양수의 eigenvalue 를 하나 이상씩 가진다.

SKEW-HERMITIAN

$A^* = -A$ 를 만족하는 square matrix A를 skew-Hermitian matrix 로 정의한다.

skew-Hermitian matrix 는 normal matrix 이고, 그 eigenvalue 는 pure-imaginary(순허수) 또는 0 이다.

SKEW-SYMMETRIC MATRIX

real matrix A가 skew-Hermitian matrix 이면 A를 skew-symmetric matrix 라고 정의한다.

skew-symmetric matrix A는 다음과 같이 분해가 가능하다.

$$A = Q \begin{bmatrix} A_1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & A_p & \ddots & & \vdots \\ \vdots & & \ddots & 0 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 0 \end{bmatrix} Q^*$$

$$A_i = \begin{bmatrix} 0 & -\lambda_i \\ \lambda_i & 0 \end{bmatrix}$$

Q는 real orthogonal matrix 이고, $\lambda_i \in \mathbb{R}$ 이다.

4 SINGULAR VALUE DECOMPOSITION

singular value decomposition 은 m×n matrix A를 다음 조건을 만족하는 m×m unitary matrix V, m×n rectangular diagonal matrix Σ, n×n unitary matrix W 로 분해한다.

$$A = V\Sigma W^*$$

$m \geq n$일 때

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \sigma_n \\ \vdots & & & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}$$

이다. $\sigma_1, \sigma_2, \cdots, \sigma_n$를 matrix A 의 singular value 라고 부른다.

또, $m \leq n$ 일 때

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \sigma_2 & \ddots & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & \vdots \\ 0 & \cdots & 0 & \sigma_m & 0 & \cdots & 0 \end{bmatrix}$$

이다.

Singular value 는 0 이상의 실수여야 한다.

만약, A 가 real matrix 라면, V, W 도 real orthogonal matrix 로 분해할 수 있다,

singular value decomposition 을 하는 방법을 알아보자.

먼저 A^*A를 eigen decomposition 한다. A^*A 는 positive semidefinite matrix 이므로

$$D \in \mathbb{R}^{r\times r}$$

$$W^*A^*AW = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

로 분해된다. D는 대각선 원소의 값이 양의 실수인 diagonal matrix 이다.

여기서, $r = \text{rank}(A^*A) \leq n$ 이고, $r = \text{rank}(A) \leq m$ 도 성립한다. (증명은 나중에..)

$W = [W_1 \quad W_2]$로 partition 한 후 위 식에 대입하면,

$$W_1 \in \mathbb{C}^{n\times r}, W_2 \in \mathbb{C}^{n\times(n-r)}$$

$$\begin{bmatrix} W_1{}^* \\ W_2{}^* \end{bmatrix} A^*A[W_1 \quad W_2] = \begin{bmatrix} W_1{}^*A^*AW_1 & W_1{}^*A^*AW_2 \\ W_2{}^*A^*AW_1 & W_2{}^*A^*AW_2 \end{bmatrix} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

(식 4-1)

여기에서, $D^{\frac{1}{2}}$ 에 0 으로 된 column vector 또는 row vector 를 추가하여 m×n matrix Σ 를 구한다.

$$\Sigma = \begin{bmatrix} D^{\frac{1}{2}} & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{bmatrix}$$

$A\,W_1D^{-\frac{1}{2}}$ 를 QR decomposition 해서 , unitray matrix V 를 구한다.

$$A\,W_1D^{-\frac{1}{2}} = V\begin{bmatrix} R \\ 0_{m-r,r} \end{bmatrix}$$

$$V \in \mathbb{C}^{m\times m}, R \in \mathbb{R}^{r\times r}$$

(식 4-2)

여기서 $R = I_r$인 것을 확인해보자.

$$R^*R = [R^* \quad 0_{r,m-r}]V^*V\begin{bmatrix} R \\ 0_{m-r,r} \end{bmatrix} = \left(D^{-\frac{1}{2}}W_1{}^*A^*\right)\left(A\,W_1D^{-\frac{1}{2}}\right)$$

(식 4-1)에 의하면, $W_1{}^*A^*AW_1 = D$ 이므로 $R^*R = I_r$이고, 또한 R은 upper triangular matrix 이므로 $R = I_r$ 이어야 한다.

이제, $V\,\Sigma\,W^* = A$ 인지 확인해 보자.

$$V\,\Sigma\,W^* = V\begin{bmatrix} D^{\frac{1}{2}} & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{bmatrix} W^*$$

$$= \left(V\begin{bmatrix} I_r \\ 0_{m-r,r} \end{bmatrix}\right)\left(\begin{bmatrix} D^{\frac{1}{2}} & 0_{r,n-r} \end{bmatrix}\begin{bmatrix} W_1{}^* \\ W_2{}^* \end{bmatrix}\right)$$

(식 4-2)에 의해 $A\,W_1D^{-\frac{1}{2}} = V\begin{bmatrix} I_r \\ 0_{m-r,r} \end{bmatrix}$ 이므로

$$= \left(A\,W_1D^{-\frac{1}{2}}\right)\left(D^{\frac{1}{2}}W_1{}^*\right)$$

$$= A\,W_1W_1{}^* = A\left([W_1 \quad W_2]\begin{bmatrix} W_1{}^* \\ W_2{}^* \end{bmatrix} - W_2W_2{}^*\right)$$

$$= A(I - W_2W_2{}^*)$$

그런데, (식 4-1) 에 의하면 $W_2{}^*A^*AW_2 = 0$ 이므로, $AW_2 = 0$ 이어야 한다. 그러므로

$$= A - (AW_2)W_2{}^* = A$$

RANK, RANGE, NULL SPACE, NONSINGULAR

m×n matrix A 의 rank 가 r 이라고 하면, A 는 다음과 같이 singular decomposition 될 수 있다.

$$A = [\mathbf{v_1} \quad \mathbf{v_2} \quad \cdots \quad \mathbf{v_m}] \begin{bmatrix} \sigma_1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \sigma_2 & \ddots & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & \ddots & \sigma_r & \ddots & & \vdots \\ \vdots & & & \ddots & 0 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{bmatrix} [\mathbf{w_1} \quad \mathbf{w_2} \quad \cdots \quad \mathbf{w_n}]^*$$

$\mathrm{rank}(A) = \mathrm{rank}(A^*) = \mathrm{rank}(A^*A) = \mathrm{rank}(AA^*)$ 임을 쉽게 알 수 있다.

matrix A 의 range 는 $\mathrm{span}\{\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_r}\}$ 이다.

또 A 의 null space 는 $\mathrm{span}\{\mathbf{w_{r+1}}, \mathbf{w_{r+2}}, \cdots, \mathbf{w_n}\}$ 이다.

$r = n$ 이면, A는 nonsingular matrix 이다.

모든 l×m nonsingular matrix B 에 대하여, $\mathrm{rank}(BA) = \mathrm{rank}(A)$이다.

BA 의 range 인 $\mathrm{span}\{B\mathbf{v_1}, B\mathbf{v_2}, \cdots, B\mathbf{v_r}\}$ 의 dimension 이 r 이기 때문이다.

또, 모든 n×n invertible matrix C 에 대하여, C^*는 nonsingular matrix 이므로

$$\mathrm{rank}(AC) = \mathrm{rank}((AC)^*) = \mathrm{rank}(C^*A^*) = \mathrm{rank}(A^*) = \mathrm{rank}(A)$$

이다.

SINGULAR VALUE

m×n matrix A 의 singular values 중 가장 큰 값을 $\sigma_{\max}$라고 하고, 가장 작은 값을 $\sigma_{\min}$라고 하자.

벡터 $\mathbf{x}$ 에 대하여

$$\sigma_{\max} = \max_{x \neq 0} \frac{|A\mathbf{x}|}{|\mathbf{x}|}$$

임을 증명해보자.

$$\frac{|A\mathbf{x}|^2}{|\mathbf{x}|^2} = \frac{\mathbf{x}^* A^* A\mathbf{x}}{\mathbf{x}^*\mathbf{x}} = \frac{\mathbf{x}^*(W\,\Sigma\,V^*)(V\,\Sigma\,W^*)\mathbf{x}}{\mathbf{x}^*\mathbf{x}} = \frac{\mathbf{x}^* W\,\Sigma^2\,W^*\mathbf{x}}{\mathbf{x}^*\mathbf{x}}$$

$W^*\mathbf{x}$를 $\boldsymbol{y}$로 표기하면

$$= \frac{\mathbf{y}^*\,\Sigma^2\,\boldsymbol{y}}{\mathbf{y}^* W^* W\mathbf{y}} = \frac{\mathbf{y}^*\,\Sigma^2\,\boldsymbol{y}}{\mathbf{y}^*\mathbf{y}} = \frac{|\Sigma\,\mathbf{y}|^2}{|\mathbf{y}|^2}$$

마찬가지로

$$\sigma_{\min} = \min_{x \neq 0} \frac{|A\mathbf{x}|}{|\mathbf{x}|}$$

여기서 $\mathbf{x}$ 가 eigenvector 일 수도 있으므로,

$$\sigma_{\min} \le \min_i |\lambda_i| \le \max_i |\lambda_i| \le \sigma_{\max}$$

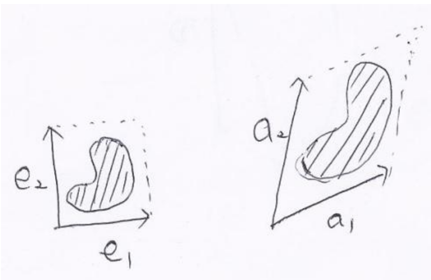
도 성립한다.

DETERMINANT

Determinant 의 기하학적 의미를 알아보자.

2×2 matrix A 의 column vector 가 $\mathbf{a_1}, \mathbf{a_2}$라고 하자

$$A = [\mathbf{a_1} \quad \mathbf{a_2}]$$



그러면, $\mathbf{a_1} = A\mathbf{e_1}$, $\mathbf{a_2} = A\mathbf{e_2}$ 를 각 변으로 하는 평행사변형의 면적이 $|\det(A)|$ 이다.

다르게 표현하면, 어떤 도형이 matrix A 에 의하여 변환되었을 때, 그 변환된 도형의 면적이 확대되는 비율이다. 3×3 matrix 의 경우에는 부피가 확대되는 비율이다.

square matrix A, B 에 대하여

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

인 법칙이 있는데, 이런 기하학적 해석에 부합된다.

Square matrix A 에 대하여,

$$\det(A) = \prod_i \lambda_i$$

$$|\det(A)| = \prod_i \sigma_i$$

가 성립한다.

$A = V\Sigma W^*$ 가 real square matrix 이고 $\det(A) < 0$ 이면, V, W 중 하나는 왼손 좌표계와 오른손 좌표계 간에 변환하는 reflection matrix 이고, 나머지 하나는 rotation matrix 이다.

5 JORDAN DECOMPOSITION

Chapter 3 에서 배운 것처럼 모든 square matrix 에 대하여 eigen decomposition 이 가능한 것은 아니다.

하지만, diagonal matrix 대신 Jordan canonical form J 으로 바꾸는 것은 항상 가능하다.

$$A = S\,J\,S^{-1}$$

우선 Jordan canonical form 이 무엇인지부터 알아보자. 아래와 같이 생긴 k×k matrix $J_k(\lambda)$ 을 Jordan block 이라고 한다.

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix}$$

n×n Jordan matrix J 는 아래와 같이 대각선이 Jordan block 으로 이루어진 행렬이다.

$$n_1 + n_2 + \cdots n_k = n$$

$$J = \begin{bmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{n_k}(\lambda_k) \end{bmatrix}$$

Jordan matrix 를 Jordan canonical form 이라고 부르기도 한다. 모든 square matrix 는 Jordan matrix 로 변환될 수 있다. 변환해 봐서, Jordan matrix 가 diagonal matrix 이면 eigen decomposition 이 가능한 matrix 이고, 아니면 eigen decomposition 이 불가능한 matrix 이다.

Jordan matrix 를 구하는 과정이 복잡하므로, 꼭 알 필요까지는 없지만 그 결과는 알아두면 matrix 를 이해하는 데 도움이 된다. 꼭 알고 싶은 분을 위해 설명해 보겠다.

square matrix A를 Jordan matrix 로 변환하는 과정은 크게 세 단계로 이루어진다.

STEP 1.

첫 번째 단계는 Schur decomposition 을 하는 것이다.

unitary matrix Q 를 이용해 upper triangle matrix T 를 구한다.

$$Q^*AQ = T$$

T 의 대각선 원소의 순서를 정할 때, 동일한 eigenvalue 는 대각선 방향으로 인접하도록 배치한다. 아래 형식으로 변형된다.

$$T = \begin{bmatrix} T_{n_1}(\lambda_1) & * & \cdots & * \\ 0 & T_{n_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & T_{n_k}(\lambda_k) \end{bmatrix}$$

$$T_k(\lambda) = \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda \end{bmatrix} \in \mathbb{C}^{k\times k}$$

$\lambda_1, \lambda_2, \cdots, \lambda_k$ 는 모두 서로 다른 값을 가져야 한다.

STEP 2.

두 번째 단계는 $T_{n_i}(\lambda_i)$이외의 원소들의 값을 0 으로 바꾼다.

두 번째 단계가 끝나면 아래 형식으로 바뀐다.

$$\begin{bmatrix} T_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & T_{n_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & T_{n_k}(\lambda_k) \end{bmatrix}$$

p 행 q 열의 원소를 0 으로 바꾸기 위해서, 적당히 설정된 invertible matrix 를 $E_{p,q}$이용해서,

T 대신 $E_{p,q}{}^{-1}\, T\, E_{p,q}$로 변형하는 작업을 한다.

$$T = \begin{bmatrix} t_{1,1} & \cdots & \cdots & t_{1,p} & \cdots & t_{1,q} & \cdots & \cdots & t_{1,n} \\ 0 & \ddots & & t_{2,p} & & t_{2,q} & & & \vdots \\ \vdots & \ddots & \ddots & \vdots & & \vdots & & & \vdots \\ \vdots & & \ddots & t_{p,p} & \cdots & t_{p,q} & t_{p,q+1} & \cdots & t_{p,n} \\ \vdots & & & \ddots & \ddots & \vdots & & & \vdots \\ \vdots & & & & \ddots & t_{q,q} & t_{q,q+1} & \cdots & t_{q,n} \\ \vdots & & & & & \ddots & \ddots & & \vdots \\ \vdots & & & & & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & t_{n,n} \end{bmatrix}$$

라고 표시하고,

$$E_{p,q} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & & & \vdots \\ \vdots & & \ddots & 1 & \ddots & \alpha & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & & & \ddots & 1 & \ddots & & \vdots \\ \vdots & & & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{bmatrix}$$

로 정의하자. $E_{p,q}$의 p 행 q 열 원소의 값은 α 이다.

$$
E_{p,q}{}^{-1} = \begin{bmatrix}
1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & \ddots & \ddots & & & & & & \vdots \\
\vdots & \ddots & \ddots & \ddots & & & & & \vdots \\
\vdots & & \ddots & 1 & \ddots & -\alpha & & & \vdots \\
\vdots & & & \ddots & \ddots & \ddots & & & \vdots \\
\vdots & & & & \ddots & 1 & \ddots & & \vdots \\
\vdots & & & & & \ddots & \ddots & \ddots & \vdots \\
\vdots & & & & & & \ddots & \ddots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1
\end{bmatrix}
$$

$$
E_{p,q}{}^{-1}\, T\, E_{p,q} = \begin{bmatrix}
t_{1,1} & \cdots & \cdots & t_{1,p} & \cdots & t_{1,q} + \boldsymbol{\alpha}\, \mathbf{t_{1,p}} & \cdots & \cdots & t_{1,n} \\
0 & \ddots & & t_{2,p} & & t_{2,q} + \boldsymbol{\alpha}\, \mathbf{t_{2,p}} & & & \vdots \\
\vdots & \ddots & \ddots & \vdots & & \vdots & & & \vdots \\
\vdots & & \ddots & t_{p,p} & \cdots & t_{p,q} + \boldsymbol{\alpha}\, \mathbf{t_{p,p}} - \boldsymbol{\alpha}\, \mathbf{t_{q,q}} & t_{p,q+1} - \boldsymbol{\alpha}\, \mathbf{t_{q,q+1}} & \cdots & t_{p,n} - \boldsymbol{\alpha}\, \mathbf{t_{q,n}} \\
\vdots & & & \ddots & \ddots & \vdots & & & \vdots \\
\vdots & & & & \ddots & t_{q,q} & t_{q,q+1} & \cdots & t_{q,n} \\
\vdots & & & & & \ddots & \ddots & & \vdots \\
\vdots & & & & & & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & t_{n,n}
\end{bmatrix}
$$

그러므로,

$$
\alpha = \frac{-t_{p,q}}{t_{p,p} - t_{q,q}}
$$

로 정하면 T 의 p 행 q 열의 값을 0 으로 바꿀 수 있다.

그리고, 이 변환에 의하면 p 행 q 열의 위쪽과 오른쪽 원소들의 값만 바뀌므로, (p, q) 를 다음 순서대로 반복 적용한다.

$$(n_1, n_1 + 1), (n_1 - 1, n_1 + 1), \cdots, (1, n_1 + 1)$$

$$(n_1, n_1 + 2), (n_1 - 1, n_1 + 2), \cdots, (1, n_1 + 2)$$

$$\vdots$$

$$(n_1, n_1 + n_2), (n_1 - 1, n_1 + n_2), \cdots, (1, n_1 + n_2)$$

$$(n_1 + n_2, n_1 + n_2 + 1), (n_1 + n_2 - 1, n_1 + n_2 + 1), \cdots, (1, n_1 + n_2 + 1)$$

$$(n_1 + n_2, n_1 + n_2 + 2), (n_1 + n_2 - 1, n_1 + n_2 + 2), \cdots, (1, n_1 + n_2 + 2)$$

$$\vdots$$

$$(n_1 + n_2, n_1 + n_2 + n_3), (n_1 + n_2 - 1, n_1 + n_2 + n_3), \cdots, (1, n_1 + n_2 + n_3)$$

$$\vdots$$

$$\vdots$$

$$\vdots$$

$$\vdots$$

$$(n_1+\cdots+n_{k-1}, n_1\cdots+n_{k-1}+1), (n_1+\cdots+n_{k-1}-1, n_1+\cdots+n_{k-1}+1), \cdots, (1, n_1+\cdots+n_{k-1}+1)$$

$$(n_1+\cdots+n_{k-1}, n_1+\cdots+n_{k-1}+2), (n_1+\cdots+n_{k-1}-1, n_1+\cdots+n_{k-1}+2), \cdots, (1, n_1+\cdots+n_{k-1}+2)$$

$$\vdots$$

$$(n_1\cdots+n_{k-1}, n_1+\cdots+n_{k-1}+n_k), (n_1+\cdots+n_{k-1}-1, n_1+\cdots+n_{k-1}+n_k), \cdots, (1, n_1+\cdots+n_{k-1}+n_k)$$

STEP 3.

이제 세 번째 단계이다.

우선 Jordan block 의 성질에 대해 정리해 보자.

$$J_k(0)^T J_k(0) = \begin{bmatrix} 0 & 0 \\ 0 & I_{k-1} \end{bmatrix}$$

$$\left(I - J_k(0)^T J_k(0)\right) = \mathbf{e_1}\,\mathbf{e_1}^T$$

(식 5-1)

I_{k-1} 는 k-1 행 k-1 열 identity matrix 이다.

$\mathbf{e_i}$ 는 i 번째 standard unit vector 이다.

$$J_k(0)^k = 0$$

(식 5-2)

$$J_k(0)\mathbf{e_{i+1}} = \mathbf{e_i}$$

$$i = 1,2,\cdots,k-1$$

(식 5-3)

자 이제 아래와 같이 eigenvalue 가 모두 0 인 n×n upper triangular matrix A 를 Jordan matrix 으로 변환해 보자.

$$A = \begin{bmatrix} 0 & * & \cdots & * \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & * \\ 0 & \cdots & \cdots & 0 \end{bmatrix}$$

$$\begin{bmatrix} J_{m_1}(0) & 0 & \cdots & 0 \\ 0 & J_{m_2}(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{m_k}(0) \end{bmatrix}$$

$$m_1 + m_2 \cdots + m_k = n$$

이 때, $m_1 \geq m_2 \cdots \geq m_k$인 조건도 만족시킬 수 있다.

n = 1 인 경우에는 이미 A가 Jordan matrix 이다.

n > 2 인 경우에는

$$\mathbf{a} \in \mathbb{C}^{n-1}, A_1 \in \mathbb{C}^{(n-1)\times(n-1)}$$

$$A = \begin{bmatrix} 0 & \mathbf{a}^T \\ 0 & A_1 \end{bmatrix}$$

으로 분할될 수 있다. 게다가 A_1는 재귀적으로 Jordan matrix 로 변환될 수 있다.

$$S_1{}^{-1}A_1S_1 = \begin{bmatrix} J_{n_1}(0) & * & \cdots & * \\ 0 & J_{n_2}(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & J_{n_k}(0) \end{bmatrix}$$

$$n_1 + n_2 \cdots + n_k = n - 1$$

$$n_1 \geq n_2 \cdots \geq n_k$$

A 를 변환하면

$$\begin{bmatrix} 1 & 0 \\ 0 & S_1{}^{-1} \end{bmatrix} A \begin{bmatrix} 1 & 0 \\ 0 & S_1 \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{a}^T S_1 \\ 0 & S_1{}^{-1}A_1S_1 \end{bmatrix}$$

$K = \begin{bmatrix} J_{n_2}(0) & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_{n_k}(0) \end{bmatrix}$ 로 표기하면

$$\mathbf{a_1} \in \mathbb{C}^{n_1}, \mathbf{a_2} \in \mathbb{C}^{n_2\cdots+n_k}, K \in \mathbb{C}^{(n_2\cdots+n_k)x(n_2\cdots+n_k)})$$

$$\begin{bmatrix} 1 & 0 \\ 0 & S_1{}^{-1} \end{bmatrix} A \begin{bmatrix} 1 & 0 \\ 0 & S_1 \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{a_1}^T & \mathbf{a_2}^T \\ 0 & J_{n_1}(0) & 0 \\ 0 & 0 & K \end{bmatrix}$$

계속해서 $\mathbf{a_1}^T$의 원소를 제거하는 변환을 한다.

$$\begin{bmatrix} 1 & -\mathbf{a_1}^{\mathrm{T}} \mathrm{J}_{n_1}(0)^{\mathrm{T}} & 0 \\ 0 & \mathrm{I} & 0 \\ 0 & 0 & \mathrm{I} \end{bmatrix} \begin{bmatrix} 0 & \mathbf{a_1}^{\mathrm{T}} & \mathbf{a_2}^{\mathrm{T}} \\ 0 & \mathrm{J}_{n_1}(0) & 0 \\ 0 & 0 & \mathrm{K} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{a_1}^{\mathrm{T}} \mathrm{J}_{n_1}(0)^{\mathrm{T}} & 0 \\ 0 & \mathrm{I} & 0 \\ 0 & 0 & \mathrm{I} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \mathbf{a_1}^{\mathrm{T}} \left(\mathrm{I} - \mathrm{J}_{n_1}(0)^{\mathrm{T}} \mathrm{J}_{n_1}(0) \right) & \mathbf{a_2}^{\mathrm{T}} \\ 0 & \mathrm{J}_{n_1}(0) & 0 \\ 0 & 0 & \mathrm{K} \end{bmatrix}$$

(식 5-1)을 적용하면

$$= \begin{bmatrix} 0 & (\mathbf{a_1}^{\mathrm{T}} \mathbf{e_1})\, \mathbf{e_1}^{\mathrm{T}} & \mathbf{a_2}^{\mathrm{T}} \\ 0 & \mathrm{J}_{n_1}(0) & 0 \\ 0 & 0 & \mathrm{K} \end{bmatrix}$$

(식 5-4)

$(\mathbf{a_1}^{\mathrm{T}} \mathbf{e_1})$ 의 값에 따라 처리방법이 달라진다.

$(\mathbf{a_1}^{\mathrm{T}} \mathbf{e_1}) \neq 0$ 이면

$$\begin{bmatrix} (\mathbf{a_1}^{\mathrm{T}} \mathbf{e_1})^{-1} & 0 & 0 \\ 0 & \mathrm{I} & 0 \\ 0 & 0 & (\mathbf{a_1}^{\mathrm{T}} \mathbf{e_1})^{-1}\mathrm{I} \end{bmatrix} \begin{bmatrix} 0 & (\mathbf{a_1}^{\mathrm{T}} \mathbf{e_1})\, \mathbf{e_1}^{\mathrm{T}} & \mathbf{a_2}^{\mathrm{T}} \\ 0 & \mathrm{J}_{n_1}(0) & 0 \\ 0 & 0 & \mathrm{K} \end{bmatrix} \begin{bmatrix} (\mathbf{a_1}^{\mathrm{T}} \mathbf{e_1}) & 0 & 0 \\ 0 & \mathrm{I} & 0 \\ 0 & 0 & (\mathbf{a_1}^{\mathrm{T}} \mathbf{e_1})\mathrm{I} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \mathbf{e_1}^{\mathrm{T}} & \mathbf{a_2}^{\mathrm{T}} \\ 0 & \mathrm{J}_{n_1}(0) & 0 \\ 0 & 0 & \mathrm{K} \end{bmatrix}$$

$\begin{bmatrix} 0 & \mathbf{e_1}^{\mathrm{T}} \\ 0 & \mathrm{J}_{n_1}(0) \end{bmatrix}$는 $\mathrm{J}_{n_1+1}(0)$ 이므로

$$= \begin{bmatrix} \mathrm{J}_{n_1+1}(0) & \mathbf{e_1}\, \mathbf{a_2}^{\mathrm{T}} \\ 0 & \mathrm{K} \end{bmatrix}$$

이번엔 $\mathbf{e_1}\, \mathbf{a_2}^{\mathrm{T}}$를 제거한다.

$$\begin{bmatrix} \mathrm{I} & \mathbf{e_2}\, \mathbf{a_2}^{\mathrm{T}} \\ 0 & \mathrm{I} \end{bmatrix} \begin{bmatrix} \mathrm{J}_{n_1+1}(0) & \mathbf{e_1}\, \mathbf{a_2}^{\mathrm{T}} \\ 0 & \mathrm{K} \end{bmatrix} \begin{bmatrix} \mathrm{I} & -\mathbf{e_2}\, \mathbf{a_2}^{\mathrm{T}} \\ 0 & \mathrm{I} \end{bmatrix}$$

$$= \begin{bmatrix} \mathrm{J}_{n_1+1}(0) & -\mathrm{J}_{n_1+1}(0)\mathbf{e_2}\, \mathbf{a_2}^{\mathrm{T}} + \mathbf{e_1}\, \mathbf{a_2}^{\mathrm{T}} + \mathbf{e_2}\, \mathbf{a_2}^{\mathrm{T}}\mathrm{K} \\ 0 & \mathrm{K} \end{bmatrix}$$

(식 5-3)을 이용하면

$$= \begin{bmatrix} \mathrm{J}_{n_1+1}(0) & \mathbf{e_2}\, \mathbf{a_2}^{\mathrm{T}}\mathrm{K} \\ 0 & \mathrm{K} \end{bmatrix}$$

$\mathrm{i} = 2,3,\cdots,n_1$를 반복하면

$$\begin{bmatrix} \mathrm{I} & \mathbf{e_{i+1}}\, \mathbf{a_2}^{\mathrm{T}}\mathrm{K}^{i-1} \\ 0 & \mathrm{I} \end{bmatrix} \begin{bmatrix} \mathrm{J}_{n_1+1}(0) & \mathbf{e_i}\, \mathbf{a_2}^{\mathrm{T}}\mathrm{K}^{i-1} \\ 0 & \mathrm{K} \end{bmatrix} \begin{bmatrix} \mathrm{I} & -\mathbf{e_{i+1}}\, \mathbf{a_2}^{\mathrm{T}}\mathrm{K}^{i-1} \\ 0 & \mathrm{I} \end{bmatrix}$$

$$= \begin{bmatrix} J_{n_1+1}(0) & \mathbf{e_{i+1}}\,\mathbf{a_2}^T K^i \\ 0 & K \end{bmatrix}$$

(식 5-2)에 의해 $K^{n_1} = 0$ 이므로, 위 식은 $\begin{bmatrix} J_{n_1+1}(0) & 0 \\ 0 & K \end{bmatrix}$ 으로 변환될 수 있다.

(식 5-4)에서 $(\mathbf{a_1}^T\mathbf{e_1}) = 0$ 이면 더 간단하다..

$$\begin{bmatrix} 0 & 0 & \mathbf{a_2}^T \\ 0 & J_{n_1}(0) & 0 \\ 0 & 0 & K \end{bmatrix}$$

위 식은

$$\begin{bmatrix} 0 & I_{n_1} & \mathbf{0} \\ 1 & 0 & 0 \\ 0 & 0 & I_{n_2\cdots+n_k} \end{bmatrix}\begin{bmatrix} 0 & 0 & \mathbf{a_2}^T \\ 0 & J_{n_1}(0) & 0 \\ 0 & 0 & K \end{bmatrix}\begin{bmatrix} 0 & 1 & \mathbf{0} \\ I_{n_1} & 0 & 0 \\ 0 & 0 & I_{n_2\cdots+n_k} \end{bmatrix}$$
$$= \begin{bmatrix} J_{n_1}(0) & 0 & 0 \\ 0 & 0 & \mathbf{a_2}^T \\ 0 & 0 & K \end{bmatrix}$$

으로 변환되고,

$\begin{bmatrix} 0 & \mathbf{a_2}^T \\ 0 & K \end{bmatrix}$는 Jordan matrix A_2으로 변환 가능하므로,

$$\begin{bmatrix} I_{n_1} & 0 \\ 0 & {S_2}^{-1} \end{bmatrix}\begin{bmatrix} J_{n_1}(0) & 0 & 0 \\ 0 & 0 & \mathbf{a_2}^T \\ 0 & 0 & K \end{bmatrix}\begin{bmatrix} I_{n_1} & 0 \\ 0 & S_2 \end{bmatrix} = \begin{bmatrix} J_{n_1}(0) & 0 \\ 0 & A_2 \end{bmatrix}$$

A가

$$A = \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda \end{bmatrix}$$

일 때, Jordan matrix 로 변환하는 방법으로 확장해 보자.

$A - \lambda I$를 Jordan matrix J로 J변환된다면

$$A - \lambda I = S\,J\,S^{-1}$$
$$A = S\,J\,S^{-1} + \lambda I$$
$$= S(J + \lambda I)S^{-1}$$

A는 Jordan matrix $J + \lambda I$ 로 변환된다.

UNIQUENESS

여기까지 Jordan matrix 를 구하는 방법을 알아보았다.

마지막으로 square matrix A 의 eigenvalue 의 순서가 정해지면(엄밀하게 말해서 $J_k(\lambda_i)$의 순서가 정해지면) Jordan matrix 가 유일하게 정해진다는 것을 증명하겠다.

$$J_k(\lambda)^m = \begin{bmatrix} \binom{m}{0}\lambda^m & \binom{m}{1}\lambda^{m-1} & \cdots & \binom{m}{m} & \cdots & 0 \\ 0 & \binom{m}{0}\lambda^m & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \binom{m}{m} \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \binom{m}{0}\lambda^m & \binom{m}{1}\lambda^{m-1} \\ 0 & \cdots & \cdots & \cdots & 0 & \binom{m}{0}\lambda^m \end{bmatrix}$$

$J_k(\lambda)^m$의 rank 는 $\lambda \neq 0$ 이면 k 이다. $\lambda = 0$ 이고 $m \geq k$ 이면 0 이다. $\lambda = 0$ 이고 $m < k$ 이면 $k - m$ 이다.

그런데, A 의 Jordan matrix 가 $J_k(\lambda_i)$으로 이루어졌다고 하면, $(A - \lambda I)^m$ 의 rank 는 $J_k(\lambda_i - \lambda)^m$의 rank 의 합과 같아야 한다.

$(A - \lambda I)^m$ 의 rank 를 $R(m, \lambda)$로 표기하면, $R(m-1, \lambda) - R(m, \lambda)$는 $k \geq m$ 과 $\lambda = \lambda_i$을 만족하는 Jordan Block $J_k(\lambda_i)$의 수이다. 그래서, $k = m$ 과 $\lambda = \lambda_i$을 만족하는 Jordan Block $J_k(\lambda_i)$의 수는

$\big(R(m-1, \lambda) - R(m, \lambda)\big) - \big(R(m, \lambda) - R(m+1, \lambda)\big)$로 구할 수 있다.

RANK, NULL SPACE, EIGENVECTOR

n×n matrix A 의 rank 가 r 이라고 하면, A 는 다음과 같이 분해될 수 있다.

$$A = [\mathbf{s_1} \quad \mathbf{s_2} \quad \cdots \quad \mathbf{s_n}] \begin{bmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & J_{n_p}(\lambda_p) & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & J_{n_{p+q}}(0) \end{bmatrix} [\mathbf{s_1} \quad \mathbf{s_2} \quad \cdots \quad \mathbf{s_n}]^{-1}$$

$$\lambda_1 \neq 0, \lambda_2 \neq 0, \cdots, \lambda_p \neq 0$$

$$n - q = r$$

matrix A 의 null space 는 $\mathrm{span}\{\mathbf{s}_{\mathbf{n_1+n_2\cdots n_p+1}}, \mathbf{s}_{\mathbf{n_1+n_2\cdots n_{p+1}+1}}, \cdots, \mathbf{s}_{\mathbf{n_1+n_2\cdots\, n_{p+q-1}+1}}\}$ 이다.

A 의 eigenvector 는 $\mathbf{s_1}, \mathbf{s}_{\mathbf{n_1+1}}, \mathbf{s}_{\mathbf{n_1+n_2+1}}, \cdots, \mathbf{s}_{\mathbf{n_1+n_2\cdots\, n_{p+q-1}+1}}$ 로 p + q 개이다.

6 NORM

VECTOR NORM

다음 조건을 만족하는 vector 를 입력으로 하고, 실수를 출력으로 하는 함수를 vector norm 이라고 정의하고, 입력 $\mathbf{x}$ 의 함수 값을 $\|\mathbf{x}\|$ 로 표시한다.

- 모든 $\mathbf{x} \in \mathbb{C}^n$ 에 대하여, $\|\mathbf{x}\| >= 0$
- 모든 $\mathbf{x} \in \mathbb{C}^n$ 에 대하여, $\|\mathbf{x}\| = 0 \leftrightarrow \mathbf{x} = 0$
- 모든 $\alpha \in \mathbb{C}, \mathbf{x} \in \mathbb{C}^n$ 에 대하여, $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$
- 모든 $\mathbf{x} \in \mathbb{C}^n, \mathbf{y} \in \mathbb{C}^n$ 에 대하여, $\|\mathbf{x} + \mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$

Vector norm 중 중요한 것을 알아보자.

$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$라고 하자.

p 가 자연수일 때, p-norm $\|\mathbf{x}\|_p$을 다음과 같이 정의한다.

$$\|\mathbf{x}\|_p = \left(\sum_i |x_i|^p\right)^{\frac{1}{p}}$$

또, ∞-norm $\|\mathbf{x}\|_\infty$은 다음과 같이 정의한다.

$$\|\mathbf{x}\|_\infty = \max_i |x_i|$$

MATRIX NORM

모든 vector space 에 대하여, norm 을 정의할 수 있다.

다음 조건을 만족하는 matrix 를 입력으로 하고, 실수를 출력으로 하는 함수를 함수를 matrix norm 이라고 정의하고, 입력 A 의 함수 값을 $\|A\|$ 로 표시한다

- 모든 $A \in \mathbb{C}^{m\times n}$ 에 대하여, $\|A\| \geq 0$
- 모든 $A \in \mathbb{C}^{m\times n}$ 에 대하여, $\|A\| = 0 \leftrightarrow A = 0$
- 모든 $\alpha \in \mathbb{C}, A \in \mathbb{C}^{m\times n}$ 에 대하여, $\|\alpha A\| = |\alpha|\|A\|$
- 모든 $A \in \mathbb{C}^{m\times n}, B \in \mathbb{C}^{m\times n}$ 에 대하여, $\|A + B\| \leq \|A\| + \|B\|$

Matrix norm 중에는 induced norm $\|A\|_p$이 있는데, 아래와 같이 정의한다.

$$\|A\|_p = \max_{\mathbf{x}\neq 0,} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p}$$

$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$라고 하자.

induced norm $\|A\|_1$를 특별히 column sum norm 이라고 부르는데,

$$\|A\|_1 = \max_{\mathbf{x}\neq 0,} \frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$$

가 성립함을 알 수 있다.

induced norm $\|A\|_\infty$ 를 row sum norm 이라고 부르는데,

$$\|A\|_\infty = \max_{\mathbf{x}\neq 0,} \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|$$

가 성립한다.

induced norm $\|A\|_2$를 spectral norm 이라고 부르는데, A의 singular value 의 최대값이 σ_{max}라고 하면,

$$\|A\|_2 = \max_{\mathbf{x}\neq 0,} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sigma_{max}$$

이다.

또, 중요한 matrix norm 중에는 Frobenius norm $\|A\|_F$이 있는데, 다음과 정의한다.

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}$$

아래 성질은 만족하는 matrix norm 을 sub-multiplicative norm 이라고 한다.

- 모든 $A \in \mathbb{C}^{n\times n}, B \in \mathbb{C}^{n\times n}$ 에 대하여, $\|AB\| \le \|A\|\|B\|$

induced norm 과 Forbenius norm 은 sub-multiplicative norm 이다. 모든 matrix norm 이 sub-multiplicative norm 은 아니다.

7 CONCLUSION

SIMILAR

square matrix A, B 와 invertible matrix S 에 다음 식이 성립할 때,

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

A 와 B는 similar 하다고 정의한다. 이 때, A 와 B는 동일한 eigenvalue 들을 가진다.

A 와 B가 similar 하면 동일한 Jordan matrix 으로 변환되고, 동일한 Jordan matrix 으로 변환되는 A 와 B는 similar 하다.

square matrix A, B 와 unitary matrix Q 에 다음 식이 성립할 때,

$$A = QBQ^*$$

A 와 B는 unitarily equivalent 하다고 정의한다. 이 때 A 와 B는 동일한 singular value 들을 가진다.

역으로, A 와 B가 동일한 singular value 들을 가지면, A 와 B는 unitarily equivalent 하다.

TRACE

square matrix $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$ 의 trace $tr(A)$를 다음과 정의한다.

$$tr(A) = \sum_{i=1}^{n} a_{ii}$$

그리고, $m \times n$ matrix $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ 와 $n \times m$ matrix $B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{m2} & \cdots & b_{nm} \end{bmatrix}$ 에 대하여,

$$tr(AB)$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ji}$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} b_{ji}$$

$$= tr(BA)$$

이므로,

$$tr(AB) = tr(BA)$$

인 법칙이 성립한다.

모든 square matrix $A \in \mathbb{C}^{n \times n}$ 는 Schur decomposition 에 의해

$$Q \in \mathbb{C}^{n \times n}, T \in \mathbb{C}^{n \times n}$$

$$A = QTQ^*$$

로 분해될 수 있으므로,

$$tr(A) = tr(QTQ^*) = tr(TQ^*Q) = tr(T)$$

이고,

$$\mathrm{tr}(A) = \sum_i \lambda_i$$

인 성질이 있다.

두 matrix 가 similar 하면, trace 가 같다.

FROBENIUS NORM

6 장에서 Frobenius norm 에 대해 배웠다.

m×n matrix A 에 대하여

$$\|A\|_F = \sqrt{\mathrm{tr}(A^* A)}$$

이 성립하는 것을 쉽게 증명할 수 있다.

A 의 singular value 가 σ_i라고 하면,

$$\|A\|_F = \sqrt{\sum_i \sigma_i{}^2}$$

도 위의 식과 singular value decomposition 을 이용해 증명할 수 있다.

두 matrix 가 unitarily equivalent 하면, Frobenius norm 은 같다.

SIMULTANEOUS DIAGONALIZATION

square matrix A와 B가 eigen decomposition 가능하고, $AB = BA$ 라면, $S^{-1}AS$ 와 $S^{-1}BS$ 를 diagonal matrix 로 만드는 invertible matrix S 가 존재한다.

[1] 1.3.12 참조

$A \in \mathbb{C}^{n\times n}$를 Jordan decomposition 하면

$$A = S_A D_A S_A{}^{-1}$$

$$n_1 + n_2 + \cdots n_k = n$$

$$D_A = \begin{bmatrix} \lambda_1 I_{n_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{n_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_k I_{n_k} \end{bmatrix}$$

$$I_{n_1} \in \mathbb{C}^{n_1\times n_1}, I_{n_2} \in \mathbb{C}^{n_2\times n_2}, \cdots, I_{n_k} \in \mathbb{C}^{n_k\times n_k}$$

이고, $\lambda_1, \lambda_2, \cdots, \lambda_k$ 는 모두 서로 다른 값을 가진다.

$S_A{}^{-1} B S_A$ 를 아래와 같이 분할해 표현할 수 있다.

$$S_A{}^{-1} B S_A = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1k} \\ B_{21} & B_{22} & \cdots & B_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ B_{k1} & B_{k2} & \cdots & B_{kk} \end{bmatrix}$$

$$B_{11} \in \mathbb{C}^{n_1\times n_1}, B_{12} \in \mathbb{C}^{n_1\times n_2}, \cdots, B_{1k} \in \mathbb{C}^{n_1\times n_k}$$

$$B_{21} \in \mathbb{C}^{n_2\times n_1}, B_{22} \in \mathbb{C}^{n_2\times n_2}, \cdots, B_{2k} \in \mathbb{C}^{n_2\times n_k}$$

$$\vdots$$

$$B_{k1} \in \mathbb{C}^{n_k\times n_1}, B_{k2} \in \mathbb{C}^{n_k\times n_2}, \cdots, B_{kk} \in \mathbb{C}^{n_k\times n_k}$$

$AB = BA$ 이면 $D_A\left(S_A{}^{-1} B S_A\right) = \left(S_A{}^{-1} B S_A\right) D_A$ 이므로

$$\begin{bmatrix} \lambda_1 B_{11} & \lambda_1 B_{12} & \cdots & \lambda_1 B_{1k} \\ \lambda_2 B_{21} & \lambda_2 B_{22} & \cdots & \lambda_2 B_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_k B_{k1} & \lambda_k B_{k2} & \cdots & \lambda_k B_{kk} \end{bmatrix} = \begin{bmatrix} \lambda_1 B_{11} & \lambda_2 B_{12} & \cdots & \lambda_k B_{1k} \\ \lambda_1 B_{21} & \lambda_2 B_{22} & \cdots & \lambda_k B_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 B_{k1} & \lambda_2 B_{k2} & \cdots & \lambda_k B_{kk} \end{bmatrix}$$

$$S_A{}^{-1} B S_A = \begin{bmatrix} B_{11} & 0 & \cdots & 0 \\ 0 & B_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & B_{kk} \end{bmatrix}$$

$B_{11}, B_{22}, \cdots, B_{kk}$ 를 eigen decomposition 하면

$$S_A{}^{-1} B S_A = \begin{bmatrix} S_{B1} D_{B1} S_{B1}{}^{-1} & 0 & \cdots & 0 \\ 0 & S_{B2} D_{B2} S_{B2}{}^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & S_{Bk} D_{Bk} S_{Bk}{}^{-1} \end{bmatrix}$$

$S_B = \begin{bmatrix} S_{B1} & 0 & \cdots & 0 \\ 0 & S_{B1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & S_{Bk} \end{bmatrix}$, $S = S_A S_B$ 로 표기하면

$$S^{-1}BS = S_B^{-1}(S_A^{-1}BS_A)S_B = \begin{bmatrix} D_{B1} & 0 & \cdots & 0 \\ 0 & D_{B2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & D_{Bk} \end{bmatrix}$$

$$S^{-1}AS = S_B^{-1}S_A^{-1}(S_A D_A S_A^{-1})S_A S_B = S_B^{-1}D_A S_B = D_A$$

이다. ■

square matrix A, B 에 대하여, Q^*AQ 와 Q^*BQ 를 upper triangular matrix 만드는 unitary matrix Q 가 존재하면, 이를 **simultaneously triangularizable** 하다고 말한다.

square matrix A, B 가 $AB = BA$ 라면, simultaneously triangularizable 하다.

λ가 matrix $A \in \mathbb{C}^{n\times n}$의 eigenvalue 일 때, $\ker(A - \lambda I) = \{v \in \mathbb{C}^{n\times 1} | (A - \lambda I)v = 0\}$라고 하자.

$$v \in \ker(A - \lambda I) \to Av = \lambda v \to ABv = BAv = \lambda Bv \to Bv \in \ker(A - \lambda I)$$

$$v \in \ker(A - \lambda I) \to Bv \in \ker(A - \lambda I)$$

matrix B의 정의구역(domain)과 공변역(codomain)을 $\ker(A - \lambda I)$의 공간으로 제한하여 생각하면, $\ker(A - \lambda I)$에 matrix B의 eigenvector 가 한 개 이상 존재해야 한다.

그러므로, $\ker(A - \lambda I)$에는 matrix A 와 matrix B에 공통인 eigenvector 를 포함한다. 그 이후는 Schur decompostioin 과 비슷하다.

APPENDIX

COMPLEX NUMBER

Complex number(복소수) λ를 2×2 matrix 로 표현할 수가 있다.

$$C : \mathbb{C} \to \mathbb{R}^{2\times 2}$$

$$C(\lambda) = \begin{bmatrix} \mathrm{Re}(\lambda) & -\mathrm{Im}(\lambda) \\ \mathrm{Im}(\lambda) & \mathrm{Re}(\lambda) \end{bmatrix}$$

복소수의 사칙연산을 이 matrix 간의 연산으로 계산할 수가 있다.

$$C(\lambda_1 + \lambda_2) = C(\lambda_1) + C(\lambda_2)$$

$$C(\lambda_1\lambda_2) = C(\lambda_1)C(\lambda_2)$$

$$C(\lambda^{-1}) = C(\lambda)^{-1}$$

예를 들어,

$$\lambda_1 \in \mathbb{C}, \lambda_2 \in \mathbb{C}$$

$$C(\lambda_1)C(\lambda_2) = \begin{bmatrix} \mathrm{Re}(\lambda_1) & -\mathrm{Im}(\lambda_1) \\ \mathrm{Im}(\lambda_1) & \mathrm{Re}(\lambda_1) \end{bmatrix} \begin{bmatrix} \mathrm{Re}(\lambda_2) & -\mathrm{Im}(\lambda_2) \\ \mathrm{Im}(\lambda_2) & \mathrm{Re}(\lambda_2) \end{bmatrix}$$

$$= \begin{bmatrix} \mathrm{Re}(\lambda_1)\mathrm{Re}(\lambda_2) - \mathrm{Im}(\lambda_1)\mathrm{Im}(\lambda_2) & -\mathrm{Re}(\lambda_1)\mathrm{Im}(\lambda_2) - \mathrm{Im}(\lambda_1)\mathrm{Re}(\lambda_2) \\ \mathrm{Im}(\lambda_1)\mathrm{Re}(\lambda_2) + \mathrm{Re}(\lambda_1)\mathrm{Im}(\lambda_2) & -\mathrm{Im}(\lambda_1)\mathrm{Im}(\lambda_2) + \mathrm{Re}(\lambda_1)\mathrm{Re}(\lambda_2) \end{bmatrix}$$

$$= \begin{bmatrix} \mathrm{Re}(\lambda_1\lambda_2) & -\mathrm{Im}(\lambda_1\lambda_2) \\ \mathrm{Im}(\lambda_1\lambda_2) & \mathrm{Re}(\lambda_1\lambda_2) \end{bmatrix} = C(\lambda_1\lambda_2)$$

또한

$$C(\bar{\lambda}) = C(\lambda)^{\mathrm{T}}$$

$$|\lambda|^2 = \det\left(C(\lambda)\right)$$

도 성립한다.

이 개념을 확장하면, $\mathrm{A} \in \mathbb{C}^{\mathrm{m} \times \mathrm{n}}$을 아래와 같이 $C(\mathrm{A}) \in \mathbb{R}^{2\mathrm{m} \times 2\mathrm{n}}$ 으로 변환 가능하다.

$$\mathrm{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$C(\mathrm{A}) = \begin{bmatrix} C(a_{11}) & C(a_{12}) & \cdots & C(a_{1n}) \\ C(a_{21}) & C(a_{22}) & \cdots & C(a_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ C(a_{m1}) & C(a_{m2}) & \cdots & C(a_{mn}) \end{bmatrix}$$

REAL SCHUR DECOMPOSITION

real square matrix A 는 아래와 같이 real orthogonal matrix Q 와 real matrix T 로 분해할 수 있다.

$$\mathrm{A} \in \mathbb{R}^{n \times n}, \mathrm{Q} \in \mathbb{R}^{n \times n}, \mathrm{T} \in \mathbb{R}^{n \times n}$$

$$\mathrm{A} = \mathrm{QTQ}^{\mathrm{T}}$$

$$T = \begin{bmatrix} \lambda_1 & * & \cdots & \cdots & \cdots & * \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \lambda_p & \ddots & & \vdots \\ \vdots & & \ddots & A_1 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & * \\ 0 & \cdots & \cdots & \cdots & 0 & A_q \end{bmatrix}$$

$$A_k \in \mathbb{R}^{2\times 2}$$

n 이 2 이상일 때에는 A의 한 eigenvalue λ와 그 eigenvector $\mathbf{x}$ 가 존재한다. 또, λ가 실수가 아니라면, eigenvalue $\bar{\lambda}$와 그에 대응하는 eigenvector $\bar{\mathbf{x}}$도 존재한다.

$$A\mathbf{x} = \lambda\mathbf{x} \to \overline{(A\mathbf{x})} = \overline{(\lambda\mathbf{x})} \to A\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}$$

그런데,

$$\begin{bmatrix} \dfrac{\lambda\mathbf{x} + \bar{\lambda}\bar{\mathbf{x}}}{2} & \dfrac{\lambda\mathbf{x} - \bar{\lambda}\bar{\mathbf{x}}}{2i} \end{bmatrix} = \begin{bmatrix} \dfrac{\mathbf{x} + \bar{\mathbf{x}}}{2} & \dfrac{\mathbf{x} - \bar{\mathbf{x}}}{2i} \end{bmatrix} \begin{bmatrix} \dfrac{\lambda + \bar{\lambda}}{2} & \dfrac{\lambda - \bar{\lambda}}{2i} \\ -\dfrac{\lambda - \bar{\lambda}}{2i} & \dfrac{\lambda + \bar{\lambda}}{2} \end{bmatrix}$$

$$[\mathrm{Re}(\lambda\mathbf{x}) \quad \mathrm{Im}(\lambda\mathbf{x})] = [\mathrm{Re}(\mathbf{x}) \quad \mathrm{Im}(\mathbf{x})] \begin{bmatrix} \mathrm{Re}(\lambda) & \mathrm{Im}(\lambda) \\ -\mathrm{Im}(\lambda) & \mathrm{Re}(\lambda) \end{bmatrix}$$

$$[\mathrm{Re}(A\mathbf{x}) \quad \mathrm{Im}(A\mathbf{x})] = [\mathrm{Re}(\mathbf{x}) \quad \mathrm{Im}(\mathbf{x})]C(\bar{\lambda})$$

가 성립된다.

여기서, $\mathbf{x}$ 와 $\bar{\mathbf{x}}$ 가 dependant 하다면, $\lambda = \bar{\lambda}$ 이므로, λ가 실수가 아니라는 조건에 위반된다.

그래서, $\mathrm{rank}([\mathbf{x} \quad \bar{\mathbf{x}}]) = \mathrm{rank}([\mathrm{Re}(\mathbf{x}) \quad \mathrm{Im}(\mathbf{x})]) = \mathrm{rank}([\mathrm{Re}(A\mathbf{x}) \quad \mathrm{Im}(A\mathbf{x})]) = 2$ 이다.

$[\mathrm{Re}(\mathbf{x}) \quad \mathrm{Im}(\mathbf{x})]$ 에 수직(orthogonal)인 벡터들을 구하기 위해, QR decompose 하면

$$U \in \mathbb{R}^{n\times n}, R \in \mathbb{R}^{2\times 2}$$

$$[\mathrm{Re}(\mathbf{x}) \quad \mathrm{Im}(\mathbf{x})] = U \begin{bmatrix} R \\ 0 \end{bmatrix}$$

U^*AU 를 계산해보면

$$U = [[\mathrm{Re}(\mathbf{x}) \quad \mathrm{Im}(\mathbf{x})]R^{-1} \quad * \quad \cdots \quad *]$$

$$U^*AU = U^*[A[\mathrm{Re}(\mathbf{x}) \quad \mathrm{Im}(\mathbf{x})]R^{-1} \quad * \quad \cdots \quad *]$$

$$= U^*[[\mathrm{Re}(A\mathbf{x}) \quad \mathrm{Im}(A\mathbf{x})]R^{-1} \quad * \quad \cdots \quad *]$$

$$= U^*[[\mathrm{Re}(\mathbf{x}) \quad \mathrm{Im}(\mathbf{x})]C(\bar{\lambda})R^{-1} \quad * \quad \cdots \quad *]$$

$$= U^*\left[U\begin{bmatrix} R \\ 0 \end{bmatrix}C(\bar{\lambda})R^{-1} \quad * \quad \cdots \quad *\right]$$

$$= \begin{bmatrix} R\,C(\bar{\lambda})\,R^{-1} & * & \cdots & * \\ 0 & * & \cdots & * \end{bmatrix}$$

$$= \begin{bmatrix} * & * & * & \cdots & * \\ * & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & * & \cdots & * \end{bmatrix}$$

U^*AU 를 아래와 같이 partition 하고

$$U^*AU = \begin{bmatrix} A_1 & * \\ 0 & A_2 \end{bmatrix}$$

$$A_1 \in \mathbb{R}^{2\times 2}, A_2 \in \mathbb{R}^{(n-2)\times(n-2)}$$

그 다음, complex 방식처럼, A_2를 재귀적으로 decompose 하면 된다.

REAL JORDAN DECOMPOSITION

아래와 같이 생긴 2k×2k matrix $C_k(\lambda)$ 을 새로운 Jordan block 으로 정의하자..

$$C_k(\lambda) = \begin{bmatrix} C(\lambda) & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \cdots & \cdots & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \cdots & \cdots & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & C(\lambda) \end{bmatrix}$$

n×n real Jordan matrix J 는 대각선이 두 종류의 Jordan block 으로 이루어진 행렬이다.

$$J = \begin{bmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & J_{n_p}(\lambda_p) & \ddots & & \vdots \\ \vdots & & \ddots & C_{n_{p+1}}(\lambda_{p+1}) & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & C_{n_{p+q}}(\lambda_{p+q}) \end{bmatrix}$$

$$n_1 + n_2 + \cdots n_p + 2(n_{p+1} + n_{p+2} + \cdots n_{p+q}) = n$$

$\lambda_1, \lambda_2, \cdots, \lambda_p$는 실수이고, $\lambda_{p+1}, \lambda_{p+2}, \cdots, \lambda_{p+q}$ 는 실수가 아닌 복소수이다.

real square matrix A 는 아래와 같이 real invertible matrix S 와 real Jordan matrix J 로 분해할 수 있다.

$$A = S\,J\,S^{-1}$$

STEP 1.

우선, real Schur decomposition 을 이용해 변환한다.

그 다음,

$$\begin{bmatrix} \lambda_1 & * & \cdots & \cdots & \cdots & * \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & \lambda_p & \ddots & & \vdots \\ \vdots & & \ddots & C(\lambda_{p+1}) & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & * \\ 0 & \cdots & \cdots & \cdots & 0 & C(\lambda_{p+q}) \end{bmatrix}$$

형태로 변환한다.

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$의 eigenvalue 가 실수가 아니라면,

$$\alpha = \frac{a+d}{2}, \beta = \frac{\sqrt{-(a-d)^2 - 4bc}}{2}$$

eigenvalue 는 $\alpha \pm \beta\, i$, eigenvector 는 $\begin{bmatrix} \frac{a-d}{2} \pm \beta\, i \\ c \end{bmatrix}$이고,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \frac{a-d}{2} & \beta \\ c & 0 \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \begin{bmatrix} \frac{a-d}{2} & \beta \\ c & 0 \end{bmatrix}^{-1}$$

이므로,

$A_{22} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, S = \begin{bmatrix} \frac{a-d}{2} & \beta \\ c & 0 \end{bmatrix}$로 표기하면

$$\begin{bmatrix} I & 0 & 0 \\ 0 & S^{-1} & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} A_{11} & A_{12}S & A_{13} \\ 0 & S^{-1}A_{22}S & S^{-1}A_{23} \\ 0 & 0 & A_{33} \end{bmatrix}$$

$$= \begin{bmatrix} A_{11} & A_{12}S & A_{13} \\ 0 & C(\alpha - \beta i) & S^{-1}A_{23} \\ 0 & 0 & A_{33} \end{bmatrix}$$

로 변환한다. 반복하면, 대각선 행렬을 $C(\lambda)$형식으로 바꿀 수 있다.

STEP 2.

두 번째 단계에서는 아래 형식으로 바꾼다.

$$\begin{bmatrix} T_{n_1}(\lambda_1) & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & T_{n_p}(\lambda_p) & \ddots & & \vdots \\ \vdots & & \ddots & T_{n_{p+1}}\left(C(\lambda_{p+1})\right) & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & T_{n_{p+q}}\left(C(\lambda_{p+q})\right) \end{bmatrix}$$

$$T_k(\lambda) = \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda \end{bmatrix} \in \mathbb{R}^{k\times k}$$

$$T_k\left(C(\lambda)\right) = \begin{bmatrix} C(\lambda) & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & C(\lambda) \end{bmatrix} \in \mathbb{R}^{2k\times 2k}$$

우선

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

$$A_{11} \in \mathbb{C}^{p\times p}, A_{12} \in \mathbb{C}^{p\times q}, A_{22} \in \mathbb{C}^{q\times q}$$

A_{11} 와 A_{22}에 공통된 eigenvalue 가 없다면, 0 이 아닌 모든 $X \in \mathbb{C}^{p\times q}$에 대하여,

$$A_{11}X - XA_{22} \neq 0$$

라는 것을 증명해 보자.

X 의 rank 를 r 이라고 할 때, singular value decomposition 하면,

$$X = V\Sigma W^*$$

$$\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$$

$$V \in \mathbb{C}^{p\times p}, \Sigma_r \in \mathbb{C}^{r\times r}, W \in \mathbb{C}^{q\times q}$$

로 분해된다.

$$A_{11}\, X = X\, A_{22}$$

$$\rightarrow A_{11}V\Sigma W^* = V\Sigma W^* A_{22}$$

$$\rightarrow V^* A_{11} V\Sigma = \Sigma W^* A_{22} W$$

$V^* A_{11} V$ 와 $W^* A_{22} W$ 를 아래와 같이 partition 하면

$$V^* A_{11} V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$$

$$V_{11} \in \mathbb{C}^{r\times r}, \qquad V_{22} \in \mathbb{C}^{(p-r)\times(p-r)}$$

$$W^* A_{22} W = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}$$

$$W_{11} \in \mathbb{C}^{r\times r}, \qquad W_{22} \in \mathbb{C}^{(q-r)\times(q-r)}$$

계속하면,

$$\begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} V_{11}\Sigma_r & 0 \\ V_{21}\Sigma_r & 0 \end{bmatrix} = \begin{bmatrix} \Sigma_r W_{11} & \Sigma_r W_{12} \\ 0 & 0 \end{bmatrix}$$

$V_{11} = \Sigma_r W_{11} \Sigma_r{}^{-1}$이므로 V_{11}과 W_{11}의 eigenvalue 는 동일하다. V_{21}과 W_{12}는 0 임을 알 수 있다. V_{11}의 eigenvalue 집합은 A_{11}의 eigenvalue 집합의 부분집합이고, W_{11}의 eigenvalue 집합은 A_{22}의 eigenvalue 집합의 부분집합이므로 모순이 된다. 증명 끝

그 다음, $\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$ 를 $\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$로 similar 하게 변환하는 방법을 알아보겠다.

$\mathbb{C}^{p\times q}$를 입력으로 하고, $\mathbb{C}^{p\times q}$를 출력으로 하는 선형변환 Ω 을 다음과 같이 정의해 보자.

$$\Omega : \mathbb{C}^{p\times q} \rightarrow \mathbb{C}^{p\times q}$$

$$\Omega(X) = A_{11}\, X - X\, A_{22}$$

Ω 를 $p \times q$ 행 $p \times q$ 열 matrix 로 표현 가능하다. $X \neq 0 \rightarrow \Omega(X) \neq 0$ 이므로 Ω는 invertible matrix 이다. (Ω의 null space 가 공집합이므로, rank 가 $(p \times q) \times (p \times q)$ 여야 한다.)

그런데,

$$\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} + A_{11}X - XA_{22} \\ 0 & A_{22} \end{bmatrix}$$

이므로,

$$\Omega(X) = -A_{12}$$

를 만족하는 X가 존재한다.

$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$ 가 real matrix 일 때도 $\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$로 변환할 수 있는 real matrix X 가 존재한다. 이런 변환을 반복하면 Step 2 를 마칠 수 있다.

STEP 3.

이제

$$\begin{bmatrix} C(\lambda) & * & \cdots & * \\ 0 & C(\lambda) & \cdots & * \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & C(\lambda) \end{bmatrix} \in \mathbb{R}^{2k \times 2k}$$

을

$$\begin{bmatrix} C(\lambda) & C(a_{1,2}) & \cdots & C(a_{1,k}) \\ 0 & C(\lambda) & \cdots & C(a_{2,k}) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & C(\lambda) \end{bmatrix} \in \mathbb{R}^{2k \times 2k}$$

형식으로 바꾼다.

$$T = \begin{bmatrix} C(\lambda) & \cdots & \cdots & T_{1,p} & \cdots & T_{1,q} & \cdots & \cdots & T_{1,k} \\ 0 & \ddots & & T_{2,p} & & T_{2,q} & & & \vdots \\ \vdots & \ddots & \ddots & \vdots & & \vdots & & & \vdots \\ \vdots & & \ddots & C(\lambda) & \cdots & T_{p,q} & T_{p,q+1} & \cdots & T_{p,k} \\ \vdots & & & \ddots & \ddots & \vdots & & & \vdots \\ \vdots & & & & \ddots & C(\lambda) & T_{q,q+1} & \cdots & T_{q,k} \\ \vdots & & & & & \ddots & \ddots & & \vdots \\ \vdots & & & & & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & C(\lambda) \end{bmatrix}$$

$$T_{i,j} \in \mathbb{R}^{2 \times 2}$$

라고 하고

$$E_{p,q} = \begin{bmatrix} I & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & & & \vdots \\ \vdots & & \ddots & I & \ddots & X & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & & & \ddots & I & \ddots & & \vdots \\ \vdots & & & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & I \end{bmatrix}$$

$$X \in \mathbb{R}^{2\times 2}$$

고 하자.

$$E_{p,q}{}^{-1}\, T\, E_{p,q}$$

$$= \begin{bmatrix} C(\lambda) & \cdots & \cdots & T_{1,p} & \cdots & T_{1,q} + \mathbf{T_{1,p}X} & \cdots & \cdots & T_{1,k} \\ 0 & \ddots & & T_{2,p} & & T_{2,q} + \mathbf{T_{2,p}X} & & & \vdots \\ \vdots & \ddots & \ddots & \vdots & & \vdots & & & \vdots \\ \vdots & & \ddots & C(\lambda) & \cdots & T_{p,q} + \boldsymbol{C(\lambda)}\,\mathbf{X} - \mathbf{X}\,\boldsymbol{C(\lambda)} & T_{p,q+1} - \mathbf{X}\,\mathbf{T_{q,q+1}} & \cdots & T_{p,n} - \mathbf{X}\,\mathbf{T_{q,k}} \\ \vdots & & & \ddots & \ddots & \vdots & & & \vdots \\ \vdots & & & & \ddots & C(\lambda) & T_{q,q+1} & \cdots & T_{q,k} \\ \vdots & & & & & \ddots & \ddots & & \vdots \\ \vdots & & & & & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & C(\lambda) \end{bmatrix}$$

여기서, $T_{p,q} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 일 때, $X = \begin{bmatrix} \frac{-b-c}{2\,\mathrm{Im}(\lambda)} & \frac{a-d}{2\,\mathrm{Im}(\lambda)} \\ 0 & 0 \end{bmatrix}$로 설정하면,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} \mathrm{Re}(\lambda) & -\mathrm{Im}(\lambda) \\ \mathrm{Im}(\lambda) & \mathrm{Re}(\lambda) \end{bmatrix} \begin{bmatrix} \frac{-b-c}{2\,\mathrm{Im}(\lambda)} & \frac{a-d}{2\,\mathrm{Im}(\lambda)} \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} \frac{-b-c}{2\,\mathrm{Im}(\lambda)} & \frac{a-d}{2\,\mathrm{Im}(\lambda)} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathrm{Re}(\lambda) & -\mathrm{Im}(\lambda) \\ \mathrm{Im}(\lambda) & \mathrm{Re}(\lambda) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{a+d}{2} & \frac{b-c}{2} \\ -\frac{b-c}{2} & \frac{a+d}{2} \end{bmatrix} = C\left(\frac{a+d}{2} - \frac{b-c}{2} i\right)$$

Complex Jordan decomposition 의 Step 2 에서 사용한 순서와 비슷하게 $E_{p,q}$를 적용하면, 모두 $C(a_{i,j})$형식으로 바꿀 수 있다.

그 다음부터는, complex Jordan decomposition 방식과 동일하다. $C(a_{i,j})$사이의 사칙연산이 complex 의 사칙연산과 동일하게 계산되므로, complex 방식과 같은 방법으로 decomposition 할 수 있다. 즉

$$\begin{bmatrix} C(\lambda) & C(a_{1,2}) & \cdots & C(a_{1,k}) \\ 0 & C(\lambda) & \cdots & C(a_{2,k}) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & C(\lambda) \end{bmatrix} \in \mathbb{R}^{2k\times 2k}$$

을

$$\begin{bmatrix} \lambda & a_{1,2} & \cdots & a_{1,k} \\ 0 & \lambda & \cdots & a_{2,k} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda \end{bmatrix} \in \mathbb{C}^{k\times k}$$

으로 간주해서 계산한다.

NORM

다음 테이블은 $\mathbf{x} \in \mathbb{C}^n$ 에 대하여, $\|\mathbf{x}\|_\alpha / \|\mathbf{x}\|_\beta$의 최대값을 나타낸다.

$\|\mathbf{x}\|_\alpha / \|\mathbf{x}\|_\beta$	$\beta = 1$	$\beta = 2$	$\beta = \infty$
$\alpha = 1$	1	$\sqrt{n}$	n
$\alpha = 2$	1	1	$\sqrt{n}$
$\alpha = \infty$	1	1	1

- $\|\mathbf{x}\|_1 / \|\mathbf{x}\|_2 : (\sum_i |x_i|)^2 \le (\sum_i 1)(\sum_i |x_i|^2)$ (Cauchy − Schwarz inequality) $ex)\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}$
- $\|\mathbf{x}\|_1 / \|\mathbf{x}\|_\infty : \sum_i |x_i| \le n \max_i(|x_i|)\ ex)\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}$
- $\|\mathbf{x}\|_2 / \|\mathbf{x}\|_1 : \sum_i |x_i|^2 \le (\sum_i |x_i|)^2\ \ ex)\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}$
- $\|\mathbf{x}\|_2 / \|\mathbf{x}\|_\infty : \sum_i |x_i|^2 \le n \max_i(|x_i|^2)\ ex)\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}$
- $\|\mathbf{x}\|_\infty / \|\mathbf{x}\|_1 : \max_i(|x_i|) \le \sum_i |x_i|\ \ ex)\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}$
- $\|\mathbf{x}\|_\infty / \|\mathbf{x}\|_2 : \max_i(|x_i|^2) \le \sum_i |x_i|^2\ \ ex)\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}$

다음 테이블은 $A \in \mathbb{C}^{m\times n}$ 에 대하여, $\|A\|_\alpha / \|A\|_\beta$의 최대값을 나타낸다. 단, $r = \mathrm{rank}(A)$

$\|A\|_\alpha / \|A\|_\beta$	$\beta = 1$	$\beta = 2$	$\beta = \infty$	$\beta = F$
$\alpha = 1$	1	$\sqrt{m}$	m	$\sqrt{m}$
$\alpha = 2$	$\sqrt{n}$	1	$\sqrt{m}$	1

$\alpha = \infty$	n	$\sqrt{\mathrm{n}}$	1	$\sqrt{\mathrm{n}}$
$\alpha = \mathrm{F}$	$\sqrt{\mathrm{n}}$	$\sqrt{\mathrm{r}}$	$\sqrt{\mathrm{m}}$	1

- $\|A\|_1/\|A\|_2 : \frac{\|Ax\|_1}{\|x\|_1} \Big/ \frac{\|Ax\|_2}{\|x\|_2} = \frac{\|Ax\|_1}{\|Ax\|_2}\frac{\|x\|_2}{\|x\|_1} \le \max_{\mathbf{y}\neq\mathbf{0}} \frac{\|y\|_1}{\|y\|_2} \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|x\|_2}{\|x\|_1}$ $ex)$ $\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}$
- $\|A\|_1/\|A\|_\infty$: similar to $\|A\|_1/\|A\|_2$ $ex)$ $\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}$
- $\|A\|_1/\|A\|_F : \max_{1\le j\le n} \left(\left(\sum_{i=1}^m |a_{ij}|\right)^2 \right) \le \sum_{j=1}^n \left(\sum_{i=1}^m |a_{ij}|\right)^2 \le \sum_{j=1}^n \left(\left(\sum_{k=1}^m 1\right) \left(\sum_{i=1}^m |a_{ij}|^2\right) \right)$

  (Cauchy − Schwarz inequality) $ex)$ $\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}$
- $\|A\|_2/\|A\|_1$: similar to $\|A\|_1/\|A\|_2$ $ex)$ $\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$
- $\|A\|_2/\|A\|_\infty$: similar to $\|A\|_1/\|A\|_2$ $ex)$ $\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}$
- $\|A\|_2/\|A\|_F : \sigma_{max}{}^2 \le \sum_i \sigma_i{}^2$ $ex)$ $\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}$
- $\|A\|_\infty/\|A\|_1$: similar to $\|A\|_1/\|A\|_2$ $ex)$ $\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$
- $\|A\|_\infty/\|A\|_2$: similar to $\|A\|_1/\|A\|_2$ $ex)$ $\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$
- $\|A\|_\infty/\|A\|_F$: similar to $\|A\|_1/\|A\|_F$ $ex)$ $\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$
- $\|A\|_F/\|A\|_1 : \sum_{j=1}^n \sum_{i=1}^m |a_{ij}|^2 \le \sum_{j=1}^n \left(\sum_{i=1}^m |a_{ij}|\right)^2 \le n \max_{1\le j\le n} \left(\left(\sum_{i=1}^m |a_{ij}|\right)^2 \right)$ $ex)$ $\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$
- $\|A\|_F/\|A\|_2 : \sum_i \sigma_i{}^2 \le r\,\sigma_{max}{}^2$ $ex)$ $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$
- $\|A\|_F/\|A\|_\infty$: similar to $\|A\|_F/\|A\|_1$ $ex)$ $\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}$

---

n×n matrix A 에 대한 함수 $\rho(A)$를 아래와 같이 정의하고, spectral radius 라고 부른다..

$$\rho(A) = \max_i |\lambda_i(A)|$$

$\lambda_i(\mathrm{A})$ 는 A 의 i 번째 eigenvalue 이다. $\mathrm{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ 라면, $\rho(\mathrm{A}) = 0 \leftrightarrow \mathrm{A} = 0$ 이 성립하지 않으므로 $\rho(\mathrm{A})$ 는 matrix norm 이 아니다.

모든 sub-multiplicative matrix norm ‖A‖에 대해

$$\rho(\mathrm{A}) \le \|\mathrm{A}\|$$

이 성립하는 것을 증명하겠다.

eigenvalue $\lambda = \rho(\mathrm{A})$에 해당하는 eigenvector 를 $\mathbf{x}$ 라고 하자. $\mathrm{X} = [\mathbf{x} \quad \mathbf{x} \quad \cdots \quad \mathbf{x}] \in \mathbb{C}^{\mathrm{n}\times\mathrm{n}}$ 이면

$$|\lambda|\|\mathrm{X}\| = \|\lambda\,\mathrm{X}\| = \|\mathrm{AX}\| \le \|\mathrm{A}\|\|\mathrm{X}\|$$

$$|\lambda| \le \|\mathrm{A}\|$$

이다. $\rho(\mathrm{A})$ 는 A 에 대한 모든 sub-multiplicative matrix norm 값의 하한(lower bound) 이다.

이번엔, $\rho(\mathrm{A})$ 가 A 에 대한 모든 sub-multiplicative matrix norm 값의 infimum (greatest lower bound) 이라는 것을 증명하자.

A 를 아래와 같이 Schur decomposition 했다면,

$$\mathrm{A} = \mathrm{QTQ}^*$$

$$\mathrm{T} = \begin{bmatrix} \lambda_1 & \mathrm{t}_{12} & \cdots & \mathrm{t}_{1\mathrm{n}} \\ 0 & \lambda_2 & \cdots & \mathrm{t}_{2\mathrm{n}} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_\mathrm{n} \end{bmatrix}$$

그리고, 실수 x 를 입력하는 n×n matrix $\mathrm{D}(x)$ 를 아래와 같이 정의하면

$$\mathrm{D}(x) = \begin{bmatrix} x^1 & 0 & \cdots & 0 \\ 0 & x^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & x^\mathrm{n} \end{bmatrix}$$

그러면,

$$\mathrm{D}(x)\,\mathrm{T}\,\mathrm{D}(x)^{-1} = \begin{bmatrix} \lambda_1 & x^{-1}\mathrm{t}_{12} & x^{-2}\mathrm{t}_{13} & \cdots & x^{-\mathrm{n}+1}\mathrm{t}_{1\mathrm{n}} \\ 0 & \lambda_2 & x^{-1}\mathrm{t}_{23} & \cdots & x^{-\mathrm{n}+2}\mathrm{t}_{2\mathrm{n}} \\ \vdots & \ddots & \lambda_3 & \cdots & x^{-\mathrm{n}+3}\mathrm{t}_{3\mathrm{n}} \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \lambda_\mathrm{n} \end{bmatrix}$$

x 를 크게 설정하면, 임의의 양의 실수 ε 에 대하여 $\|D(x)\,T\,D(x)^{-1}\|_1 < \rho(A) + \varepsilon$ 를 만족하는 x 가 존재한다.

이 큰 x 값을 사용하여, matrix norm $f(B)$를 아래처럼 정의하면

$$f(B) = \|D(x)\,Q^*BQ\,D(x)^{-1}\|_1$$

$f(A) < \rho(A) + \varepsilon$ 이라는 것을 알 수 있다.

$$\text{모든 } A \in \mathbb{C}^{n\times n} \text{ 에 대하여}, \|A\|_2 \le \sqrt{\|A\|_1 \|A\|_\infty}$$

가 성립하는 것도 증명하자.

$\rho(A^*A)$ 가 sub-multiplicative matrix norm $\|A^*A\|_1$의 하한이라는 것을 이용하면,

$$\sigma_{max}{}^2 = \rho(A^*A) \le \|A^*A\|_1 \le \|A^*\|_1 \|A\|_1 = \|A\|_\infty \|A\|_1$$

위 식은 σ_{max} 의 근사치를 간단하게 구할 때 많이 이용된다.

마지막으로, $A \in \mathbb{C}^{n\times n}$ 일 때, 어떤 $\|A\| < 1$을 만족하는 sub-multiplicative matrix norm 이 존재하면, $\lim_{k\to\infty} A^k = 0$ 이라는 것을 증명하겠다.

A 가 QTQ^* 로 Schur decomposition 되었다면, $A^k = Q\,T^k\,Q^*$이다. $\rho(A) \le \|A\| < 1$ 이므로 $\lim_{k\to\infty} T^k = 0$ 이고, $\lim_{k\to\infty} A^k = 0$ 이다.

REFERENCES

[1] Matrix Analysis, Roger A. Horn and Charles R. Johnson

행렬에 대한 고급 수학을 공부하기에 좋은 책이다. 조금 어렵다.

[2] Matrix Computations, Gene H. Golub. Charles F. Van Loan. Third Edition

행렬 계산을 컴퓨터에서 구현하는 방법에 대해서 잘 설명되어 있다. 다 읽을 필요는 없고 중간 중간 필요한 부분만 읽으면 될 듯 하다.