5장 부록

보충 설명: 5.2.2절 — 정리 5-2.1 (특이값 분해) 증명

(증명 스케치): 자세한 증명은 [1] 또는 [2] 을 참조한다.

여기서는 존재성에 대한 간략한 귀납적 증명을 설명한다. 우선 $\sigma_1 = \|A\|_2$로 정의하고, 행렬 노름의 정의에 의해 다음을 만족하는 벡터가 존재한다:

$$A\boldsymbol{v}_1 = \sigma_1 \boldsymbol{u}_1, \quad \|\boldsymbol{u}_1\| = \|\boldsymbol{v}_1\| = 1 \tag{1}$$

$\boldsymbol{v}_1$과 $\boldsymbol{u}_1$을 각각 $\mathbb{R}^n$의 정규 직교 기저 $\{\boldsymbol{v}_j\}$ 과 $\mathbb{R}^m$ 의 정규직교기저 $\{\boldsymbol{u}_j\}$로 확장한 후, 이 벡터들로 구성된 직교 행렬 V_1, U_1를 정의하자. 이때 다음이 성립한다:

$$S := U_1^T A V_1 = \begin{bmatrix} \sigma_1 & w^T \\ 0 & B \end{bmatrix} \tag{2}$$

여기서 w^T는 $(n-1)$차원 행벡터, B는 $(m-1)\times(n-1)$ 행렬이다. 여기서,

$$\left\| \begin{bmatrix} \sigma_1 & w^T \\ 0 & B \end{bmatrix} \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \right\|_2 \geq \sigma_1^2 + w^T w = (\sigma_1^2 + w^T w)^{1/2} \left\| \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \right\|_2 \tag{3}$$

을 이용하면 $\|S\|_2 \geq (\sigma_1^2 + w^T w)^{1/2}$가 된다. 이제 $\|S\|_2 = \|U_1^T A V_1\|_2 = \|A\|_2 = \sigma_1$임을 이용하면, $w = 0$임을 보일 수 있고, B는 A가 $\boldsymbol{v}_1$ 방향 외의 성분에 대해 어떻게 작용하는지를 나타낸다. 이후 B에 귀납 가정을 적용하면 $B = U_2 \Sigma_2 V_2^T$의 SVD가 존재하며, 이로부터 다음의 전체 분해가 구성된다:

$$A = U_1 \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & V_2 \end{bmatrix}^T V_1^T \tag{4}$$

이를 통해 A의 SVD 존재성을 귀납적으로 증명할 수 있다. ¶

복소 공간 경우 보충 설명

복소공간 $\mathbb{C}^n$과 내적 정의

복소공간 $\mathbb{C}^n$은 n개의 복소수(complex number) 원소를 가지는 벡터 $\boldsymbol{x}$들의 집합이다. 복소 벡터는 다음과 같이 표현된다:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \boldsymbol{x}^* = [\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n] \tag{5}$$

여기서 $x_j = a_j + ib_j$는 복소수이며, $\bar{x}_j = a_j - ib_j$는 그 켤레복소수(conjugate)이다.

복소 벡터 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$에 대해, 내적(inner product)과 노름(norm)은 다음과 같이 정의된다:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^* \boldsymbol{y} = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n \tag{6}$$

복소 행렬 A의 **켤레복소수 전치(conjugate transpose)**는 다음과 같이 정의된다:

$$A^* = \bar{A}^T, \quad (A^*)_{ij} = \bar{A}_{ji}, \quad (AB)^* = B^* A^* \tag{7}$$

> **정의 5A.1 (복소 행렬의 유형)**

복소수 정방행렬 A에 대해 다음과 같이 정의한다:

- (i) $A = A^*$이면 **에르미트(Hermitian)** 행렬이라 한다.
- (ii) $AA^* = I$이면 **유니터리(unitary)** 행렬이라 한다.
- (iii) $AA^* = A^*A$이면 **정규(normal)** 행렬이라 한다.

실수 행렬의 경우, 에르미트 행렬은 **대칭 행렬**, 유니터리 행렬은 **직교 행렬**과 동일하다.

보조정리 5A.1 (Hermitian 행렬의 성질)

복소 정방행렬 A가 에르미트(Hermitian) 행렬이면 다음이 성립한다:

- (i) 임의의 복소 벡터 $\boldsymbol{x}$에 대해 $\boldsymbol{x}^* A \boldsymbol{x}$는 **실수**이다.
- (ii) A의 모든 **고유치**는 **실수**이다.
- (iii) 서로 다른 고유치에 대응하는 **고유벡터들은 서로 직교**한다.

(증명): (i)

$$(\boldsymbol{x}^* A \boldsymbol{x})^* = \boldsymbol{x}^* A \boldsymbol{x} \Rightarrow \text{실수} \tag{8}$$

(ii) $A\boldsymbol{x} = \lambda \boldsymbol{x}$이면,

$$\boldsymbol{x}^* A \boldsymbol{x} = \lambda \boldsymbol{x}^* \boldsymbol{x} \in \mathbb{R} \Rightarrow \lambda \in \mathbb{R} \tag{9}$$

(iii)

$$
\begin{aligned}
& A\boldsymbol{x}=\lambda_1\boldsymbol{x}, \quad A\boldsymbol{y}=\lambda_2\boldsymbol{y} \quad (\lambda_1\neq\lambda_2) \\
\Rightarrow \quad & \lambda_1\boldsymbol{x}^*\boldsymbol{y}=\boldsymbol{x}^*A\boldsymbol{y}=(A\boldsymbol{x})^*\boldsymbol{y}=(\lambda_1\boldsymbol{x})^*\boldsymbol{y}=\lambda_1\boldsymbol{x}^*\boldsymbol{y} \\
& =\boldsymbol{x}^*(\lambda_2\boldsymbol{y})=\lambda_2\boldsymbol{x}^*\boldsymbol{y}\Rightarrow(\lambda_1-\lambda_2)\boldsymbol{x}^*\boldsymbol{y}=0\Rightarrow\boldsymbol{x}^*\boldsymbol{y}=0
\end{aligned} \tag{10}
$$

¶

따름정리 5A.2 (Unitary 행렬의 성질)

A가 유니터리(unitary) 행렬이면 다음이 성립한다:

- (i) 임의의 복소 벡터 $\boldsymbol{x},\boldsymbol{y}$에 대해

$$
\langle A\boldsymbol{x}, A\boldsymbol{y}\rangle=\langle\boldsymbol{x},\boldsymbol{y}\rangle, \quad \|A\boldsymbol{x}\|=\|\boldsymbol{x}\| \tag{11}
$$

- (ii) 모든 고유치는 절댓값이 1인 복소수이다: $|\lambda|=1$
- (iii) 서로 다른 고유치에 대응하는 고유벡터들은 서로 직교한다.

슈어분해(Shur Decomposition)

> **정리 5A.3 (슈어 분해, Schur Decomposition)**

임의의 복소수 정방행렬 $A\in\mathbb{C}^{n\times n}$에 대해, 다음과 같은 **슈어 분해(Schur decomposition)**가 존재한다:

$$
A=UTU^* \tag{12}
$$

여기서

- U는 **유니터리 행렬**
- T는 **상삼각 행렬(upper triangular matrix)**

(증명): A가 4×4 복소 행렬이라고 가정하자. 고유값 λ_1에 대응하는 정규화된 고유벡터 $\boldsymbol{x}_1$을 찾고, Gram–Schmidt 과정을 통해 유니터리 행렬 U_1을 구성하면,

$$
A_1=U_1^*AU_1=\begin{bmatrix}\lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & *\end{bmatrix} \tag{13}
$$

이제 $A_1(2:4,2:4)$ 부분에서 고유값 λ_2와 고유벡터 $\boldsymbol{x}_2$를 찾아, $\tilde{U}_2$를 구성한다. 이를 포함한 4×4 유니터리 행렬 U_2는 다음과 같다:

$$U_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & \tilde{U}_2 & \\ 0 & & & \end{bmatrix} \tag{14}$$

그 결과,

$$A_2 = U_2^* A_1 U_2 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix} \tag{15}$$

이 과정을 반복하여 상삼각 행렬 T를 얻을 수 있다:

$$T = U^* A U = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \lambda_3 & * \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix} \tag{16}$$

여기서 $U = U_1 U_2 U_3$는 유니터리 행렬이다. ¶

(참고): 위 과정은 고유치와 고유벡터를 알고 있다면 유한 단계 내에 종료되지만, 실제 계산에서는 반복적 알고리즘을 사용한다. 따라서 범용적인 Schur 분해는 다음과 같은 반복적 변환을 통해 구현된다:

$$\underbrace{U_j^* \dots U_2^* U_1^*}_{U^*} A \underbrace{U_1 U_2 \dots U_j}_{U} \quad \xrightarrow{j \to \infty} \quad T \tag{17}$$

여기서 각 U_j는 **하우스홀더 반사(Householder reflector)**이며, 이 과정을 통해 A는 상삼각 행렬 T로 수렴하게 된다.

복소공간 스펙트럼 정리 (Spectral Theorem in Complex Vector Spaces)

> **정리 5A.4 (복소공간 스펙트럼 정리)**

모든 정규행렬(normal matrix) $A \in \mathbb{C}^{n \times n}$는 유니터리 닮은변환(unitary similarity transformation)에 의해 대각행렬로 표현되는 스펙트럼 분해(spectral decomposition)를 갖는다:

$$A = U \Lambda U^* = \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^* + \cdots + \lambda_n \boldsymbol{u}_n \boldsymbol{u}_n^* \tag{18}$$

여기서,

- U는 유니터리 행렬 ($U^* U = I$),

- Λ는 대각행렬이며, λ_i는 A의 고유치이다.

또한, A가 **에르미트(Hermitian)** 행렬이면 Λ의 모든 대각 성분은 **실수**이다.

(증명): 정규행렬 A에 대해, Schur 분해에 따라 $A = UTU^*$라 하면, T는 상삼각 행렬이고 U는 유니터리 행렬이다. 이때 $T = U^*AU$도 정규행렬이 된다. 상삼각 행렬 T가 정규행렬일 조건은 $T^*T = TT^*$이다. 이를 항별로 비교하면 다음이 성립해야 한다:

$$(T^*T)_{ii} = \sum_{k=1}^{i} |t_{ki}|^2, \quad (TT^*)_{ii} = \sum_{k=i}^{n} |t_{ik}|^2 \tag{19}$$

$i = 1$일 때:

$$|t_{11}|^2 = |t_{11}|^2 + |t_{12}|^2 + \cdots + |t_{1n}|^2 \Rightarrow t_{1j} = 0 \text{ for } j > 1 \tag{20}$$

$i = 2$일 때도 유사하게

$$|t_{12}|^2 + |t_{22}|^2 = |t_{22}|^2 + |t_{23}|^2 + \cdots + |t_{2n}|^2 \Rightarrow t_{2j} = 0 \text{ for } j > 2 \tag{21}$$

반복하면 T는 대각행렬이 된다.

또한, A가 에르미트 행렬이면 $\Lambda^* = (U^*AU)^* = U^*A^*U = U^*AU = \Lambda$이므로 Λ는 실수 대각행렬이다.
¶

(참고): 역으로, 유니터리 닮음변환으로 대각화 가능한 행렬은 반드시 정규행렬이다. 즉, 정규행렬은 완전한 정규 직교 고유벡터 집합을 가진다는 특성이 있다.

복소공간에서의 고유치 계산 방법

정방행렬 $A \in \mathbb{C}^{n\times n}$에 대해 고유치를 수치적으로 계산하려면, 다음의 두 단계 접근법을 사용하는 것이 일반적이다. (자세한 증명은 [1] 또는 [2] 을 참조하라.)

Phase I: 상헤센버그(Hessenberg) 변환: 초기 행렬 A에 대해, 일련의 하우스홀더 반사(Householder reflectors) $U_1, U_2, \ldots, U_{n-2}$를 곱하여, A를 상헤센버그 행렬(Upper Hessenberg matrix) H로 변환한다.

$$H = U_{n-2}^* \cdots U_1^* A U_1 \cdots U_{n-2} \tag{22}$$

상헤센버그 행렬은 **주대각선 아래 한 줄만 비제로**인 구조를 가지며, QR 알고리즘 등의 반복을 위한 좋은 초기 형태가 된다.

예시 (임의의 5×5 행렬):

$$
\begin{aligned}
&\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}}_{A}
\xrightarrow{U_1^* \cdot}
\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \mathbf{0} & \times & \times & \times & \times \\ \mathbf{0} & \times & \times & \times & \times \\ \mathbf{0} & \times & \times & \times & \times \end{bmatrix}}_{U_1^* A}
\xrightarrow{\cdot U_1}
\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \times & \times & \times & \times \end{bmatrix}}_{U_1^* A U_1} \\
&\xrightarrow{U_2^* \cdot}
\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & \mathbf{0} & \times & \times & \times \\ & \mathbf{0} & \times & \times & \times \end{bmatrix}}_{U_2^* U_1^* A U_1}
\xrightarrow{\cdot U_2}
\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & \times & \times & \times \end{bmatrix}}_{U_2^* U_1^* A U_1 U_2}
\to \to
\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}}_{U_{n-2}^* \cdots U_2^* U_1^* A U_1 U_2 \cdots U_{n-2}}
\end{aligned}
\tag{23}
$$

즉,

$$
\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}}_{A}
\xrightarrow{\text{Phase I}}
\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}}_{H}
\tag{24}
$$

Phase II: QR 반복 (QR iteration): H에 대해 QR 분해 $H = QR$을 수행한 뒤,

$$
H^{(1)} = RQ, \quad H^{(2)} = Q_1^T R_1 Q_1, \quad \ldots \tag{25}
$$

을 반복하여, 수렴 시 $H^{(k)} \to T$는 A와 닮은 **상삼각 행렬**이 된다.

$$
\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}}_{A \neq A^*}
\xrightarrow{\text{Phase-I}}
\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix}}_{H}
\xrightarrow{\text{Phase-II}}
\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}}_{T}
\tag{26}
$$

에르미트 행렬의 경우

A가 에르미트 행렬이면 $A = A^*$이고, 위의 Phase I 결과 H는 **삼중 대각(tridiagonal)** 형태를 갖는다. 이 경우 고유치 계산은 더 효율적으로 이루어진다:

$$\underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}}_{A=A^*} \xrightarrow{\text{Phase I}} \underbrace{\begin{bmatrix} \times & \times & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ 0 & \times & \times & \times & 0 \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}}_{H} \xrightarrow{\text{Phase II}} \underbrace{\begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 \\ 0 & 0 & 0 & \lambda_4 & 0 \\ 0 & 0 & 0 & 0 & \lambda_5 \end{bmatrix}}_{D} \tag{27}$$

이 D는 A의 고유값을 대각선에 가지는 스펙트럼 분해 결과이다.

보충 설명: 5.2.4절 — 유사역행렬의 추가 성질

보조정리 5A.5

임의의 행렬 $A \in \mathbb{R}^{m\times n}$에 대해 다음이 성립한다.

(1) Frobenius 노름 최소화 문제의 해

$$\min_{B\in\mathbb{R}^{n\times m}} \|AB - I\|_F \tag{28}$$

이 문제의 유일한 최소 Frobenius 노름 해는 유사역행렬 A^+이다.

(2) 유사역행렬의 정사영 성질: A의 특이값 분해 $A = U\Sigma V^T$에서 $U_1 = U(:, 1:r)$, $V_1 = V(:, 1:r)$라 하면,

$$\begin{aligned} &\mathcal{N}(A) = \mathcal{N}(A^+A), \quad \mathcal{R}(A) = \mathcal{R}(AA^+) \\ &AA^+ = U_1U_1^T \quad (\text{정사영 onto } \mathcal{R}(A) \subset \mathbb{R}^m) \\ &A^+A = V_1V_1^T \quad (\text{정사영 onto } \mathcal{N}(A)^\perp \subset \mathbb{R}^n) \end{aligned} \tag{29}$$

즉, AA^+는 $\mathbb{R}^m$에서 $\mathcal{R}(A)$ 위로의 정사영이고, A^+A는 $\mathbb{R}^n$에서 $\mathcal{N}(A)^\perp$ 위로의 정사영이다.

(3) 유사역행렬의 특별한 형태

- $\text{rank}(A) = n$ (full column rank)이면,

$$A^+ = (A^TA)^{-1}A^T \tag{30}$$

- 만약 $m = n = \text{rank}(A)$ (정방행렬이고 가역)이라면,

$$A^+ = A^{-1} \tag{31}$$

(4) 유사역행렬의 기본 연산 성질

- $(A^+)^+ = A$
- $(A^+)^T = (A^T)^+$
- 일반적으로는 $(AB)^+ \neq B^+A^+$

(5) 유사역행렬과 릿지 회귀 (Ridge Regression): 다음 **정규화된 최소자승법** 문제를 고려하자:

$$\min_{\boldsymbol{x}} \|A\boldsymbol{x} - \boldsymbol{b}\|^2 + \lambda\|\boldsymbol{x}\|^2 \tag{32}$$

이 문제의 해 $\bar{\boldsymbol{x}}$는 다음 선형시스템의 해이다:

$$(A^T A + \lambda I)\bar{\boldsymbol{x}} = A^T \boldsymbol{b} \tag{33}$$

$\lambda \to 0$ 극한에서 유사역행렬은 다음과 같이 표현된다:

$$\lim_{\lambda \to 0}(A^T A + \lambda I)^{-1} A^T = \lim_{\lambda \to 0} V[(\Sigma^T \Sigma + \lambda I)^{-1} \Sigma^T] U^T = V \Sigma^+ U^T = A^+ \tag{34}$$

(6) 조건수 (Condition Number): 2-노름 기준에서 직사각 행렬 $A \in \mathbb{R}^{m \times n}$의 조건수는 다음과 같이 정의된다:

$$\kappa(A) = \|A\|_2 \cdot \|A^+\|_2 = \frac{\sigma_1}{\sigma_r} \tag{35}$$

여기서 σ_1은 가장 큰 특이값, σ_r은 0이 아닌 가장 작은 특이값이며, $\mathrm{rank}(A) = r$이다. $\kappa(A)$는 **수치적 안정성 (numerical stability)**을 평가하는 중요한 지표이다.

이러한 성질들은 유사역행렬이 단순한 최소자승 해 이상의 의미를 가지며, 정사영, 역행렬 일반화, 안정성 분석 등 다양한 분야에서 중심적 도구임을 보여준다.

각주

---

1. [Trefethen and Bau. (1997)] L. N. Trefethen and D. Bau Numerical Linear Algebra. SIAM, Philadelphia,1997. ↩ ↩

2. [Golub et al. (1995)] G. H. Golub, C. F. Van Loan, Matrix Computations. JohnsHopkins University Press, Baltimore,1995 ↩ ↩