Real Symmetric Matrix 의 대각화

(단학기 강좌용 Elementary Approach)

李 仁 碩

article version
180724

이 article 은 "單학기 선형대수학" 강좌를 위해 준비한 것이다. Real symmetric matrix 의 diagonalization 은 수학의 다양한 분야에서 이용되므로, 이를 위해 "단학기 선형대수학"의 수강생들에게 강좌의 마지막 1주일 정도를 투자하는 것도 의미 있는 계획이라고 생각된다.[1]

우리는 흔히 "선형대수학 1"에서는 1-차 연립방정식을 공부하고, "선형대수학 2"에서는 2-차 연립방정식을 다룬다고 말한다. 그래서 [I] 에서는 real symmetric matrix 의 diagonalization 을 §15.3 에서 공부한다.[2] 즉, 이 article 의 뒤에는 — 잘 드러나지는 않지만 — 2-차 연립방정식이 숨어 있다는 뜻이다.

이 article 의 목적은 $A \in \mathfrak{M}_{n,n}(\mathbb{R})$ 이 symmetric matrix 이면 A 는 언제나 diagonalizable 임을 보이는 것이다. 그리고 'elementary approach'란 (dimension 에 관한 귀납법이 아닌) **matrix size 에 관한 귀납법**을 의미한다.

이를 위해서는 우선 [I, §7.2] 까지 — 즉, (행렬의) characteristic polynomial 과 diagonalization 을 — 공부한 후, 다음

(a) $\mathbb{R}^n$ 및 $\mathbb{C}^n$ 의 dot product 와 orthonormal basis,
(b) Euclidean space 에서의 Gram-Schmidt Orthogonalization,
(c) Real orthogonal group $\mathbf{O}(n)$ 의 정의

의 준비가 필요하다. (독자들은 [I, 정리 7.3.1] 의 증명 technique 도 복습하기 바란다.)

[1] "학부 2학년 선형대수학"은 한 학기는 너무 짧고 두 학기는 좀 길다는 인식을 피할 길이 없다.
[2] [I, §15.3] 에서는 symmetric matrix보다 더 일반적인 symmetric operator를 다룬다.

제 1 절 Euclidean Space의 Orthonormal Basis

section version
150607

이제 $\mathbb{R}^n$과 $\mathbb{C}^n$의 dot product부터 시작하자. $\mathbb{R}^n$의 dot product는 너무나 익숙하다.

정의 1.1 (가) $X=(a_1,\ldots,a_n)^{\mathbf{t}}$, $Y=(b_1,\ldots,b_n)^{\mathbf{t}}\in\mathbb{R}^n$일 때,

$$\langle X,Y\rangle = X^{\mathbf{t}}\cdot Y = a_1b_1+a_2b_2+\cdots+a_nb_n$$

로 정의하고, 이를 $\mathbb{R}^n$-공간의 **dot product**라고 부른다. 그리고 dot product가 주어진 $\mathbb{R}^n$-공간을 **Euclidean space**라고 부른다.
(나) $X=(a_1,\ldots,a_n)^{\mathbf{t}}$, $Y=(b_1,\ldots,b_n)^{\mathbf{t}}\in\mathbb{C}^n$일 때,

$$\langle X,Y\rangle = X^{\mathbf{t}}\cdot \overline{Y} = a_1\overline{b_1}+a_2\overline{b_2}+\cdots+a_n\overline{b_n}$$

로 정의하고([I, 표기법 8.1.6] 참조), 이를 $\mathbb{C}^n$의 (**Hermitian**) **dot product**라고 부른다.[3]
(다) 따라서, $\mathbb{C}^n$의 Hermitian dot product를 $\mathbb{R}^n$에 **restrict**하면, $\mathbb{R}^n$의 dot product를 얻는다. 역으로, $\mathbb{C}^n$의 Hermitian dot product는 $\mathbb{R}^n$의 dot product를 $\mathbb{C}^n$으로 **extend**한 것이다.

$\mathbb{R}^n$-공간과 $\mathbb{C}^n$-공간의 차이는 '막대기(bar)'뿐.

주의 1.2 왜 $\mathbb{C}^n$의 경우에는 complex conjugation이 필요한 것일까? 만약, complex conjugation 없이 $\mathbb{C}^n$의 dot product를 정의했다면, '길이'가 0인 vector $(1+i,\,1-i)^{\mathbf{t}}$, '길이'가 허수인 vector $(i,0)^{\mathbf{t}}$ 등이 있기 때문이다.

Dot product의 기본적인 성질들은 [I, §9.1 및 §10.2]를 참조하면 된다. 이 article의 목적을 위해서는 아래 두 연습문제만으로 충분하다.

연습문제 1.3 $\mathbb{R}^n$의 dot product는 다음 성질을 갖는 것을 보여라.

(가) $\mathbb{R}^n\times\mathbb{R}^n$에서 $\mathbb{R}$로 가는 함수 $(X,Y)\mapsto\langle X,Y\rangle$는 $\mathbb{R}^n$의 bilinear form.[4]
(나) $X,Y\in\mathbb{R}^n$이면, $\langle X,Y\rangle=\langle Y,X\rangle$.
(다) $X\in\mathbb{R}^n$이면, 항상 $\langle X,X\rangle\geq 0$이고, 이때 등호가 성립할 필요충분조건은 $X=0$인 것이다.

[3] Hermitian은 수학자 C. Hermite(1822–1901)의 형용사형.
[4] 이제 미적분학에서보다 조금 고급스러운 단어를 사용하였다. [I, 정의 6.1.1] 참조. 그 덕택으로 미적분학 책에서는 넉 줄이 필요했던 것을 bilinear form이라는 말 한마디로 줄일 수 있게 되었다.

연습문제 1.4 $\mathbb{C}^n$의 Hermitian dot product는 다음 성질을 만족시키는 것을 보여라(단, $X,Y,Z\in\mathbb{C}^n$, $c\in\mathbb{C}$).

(가) $\langle X+Y,\, Z\rangle = \langle X,Z\rangle + \langle Y,Z\rangle$, $\langle cX,Y\rangle = c\,\langle X,Y\rangle$.
(나) $\langle X,\, Y+Z\rangle = \langle X,Y\rangle + \langle X,Z\rangle$, $\langle X,cY\rangle = \overline{c}\,\langle X,Y\rangle$.
(다) $\langle X,Y\rangle = \overline{\langle Y,X\rangle}$.
(라) $X\neq 0$ 이면, $\langle X,X\rangle\in\mathbb{R}$ 이고, $\langle X,X\rangle > 0$.

우리는 연습문제 1.3(다)항과 연습문제 1.4(라)항의 성질을 가르켜 dot product는 **positive definite** 라고 말한다.

Euclidean space에서 '수직'의 개념은 익숙하다. 또, 독자들은 $X\in\mathbb{R}^n$일 때, $\|X\| = \sqrt{\langle X,X\rangle}$로 정의하고, $\|X\|$를 vector X의 **길이**(크기)라고 부르는 것도 잘 알고 있을 것이다. 길이가 1인 vector는 **unit vector** 라고 부른다.

정의 1.5 $\mathbb{R}^n$에 dot product가 주어져 있을 때,

(가) $X,Y\in\mathbb{R}^n$일 때, 만약 $\langle X,Y\rangle = 0$ 이면, X와 Y는 **서로 수직**(**orthogonal** 또는 **perpendicular**)이라고 말하고, $X\perp Y$로 표기한다.[5]
(나) $\mathbb{R}^n$의 **non-zero** vector들 $X_1,\ldots,X_m$이 mutually perpendicular이면 — 즉, $[\langle X_i,X_j\rangle = 0$ for all $1\le i\neq j\le m]$이면 — $\{X_1,\ldots,X_m\}$을 $\mathbb{R}^n$의 **orthogonal subset**이라고 부른다. 이때 만약 $X_1,\ldots,X_m$이 모두 unit vector 이면, $\{X_1,\ldots,X_m\}$을 $\mathbb{R}^n$의 **orthonormal subset**이라고 부른다.
(다) $\mathbb{R}^n$의 basis $\mathfrak{B} = \{X_1,\ldots,X_n\}$이 orthonormal subset이면 $\mathfrak{B}$를 $\mathbb{R}^n$의 **orthonormal basis** 라고 부른다.

따라서, 예를 들어, $\mathbb{R}^n$의 표준기저 $\mathcal{E} = \{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$은 Euclidean space $\mathbb{R}^n$의 orthonormal basis이다.[6] 다음 연습문제는 미적분학 책에도 포함되어 있다.

연습문제 1.6 Euclidean space $\mathbb{R}^n$의 non-zero vector $X_1,\ldots,X_r$이 mutually orthogonal일 때, 즉 $[\langle X_i,X_j\rangle = 0$ for all $1\le i\neq j\le r]$일 때, 다음을 보여라.

(가) $\{X_1,\ldots,X_r\}$은 일차독립.
(나) 특별히 $r=n$일 때, $X_i' = \frac{1}{\|X_i\|}X_i$로 표기하면, $\{X_1',\ldots,X_n'\}$은 $\mathbb{R}^n$의 orthonormal basis.

[5] $X\perp Y$는 "X perp Y"로 읽는다.
[6] $\mathbb{R}^n$의 orthonormal basis의 다양한 보기는 [I, 연습문제 9.1.13] 참조.

이제 Gram-Schmidt Orthogonalization을 소개한다. 앞으로 $\mathbb{R}^n$-공간은 항상 dot product가 주어진 Euclidean space로 생각한다.

Gram-Schmidt Orthogonalization의 idea는 고등학교에서 배운 '정사영'의 개념에 포함되어 있었다. 즉, X_1, X_2, X_3가 $\mathbb{R}^n$의 non-zero vector들일 때, X_2에서 X_1-성분 $\frac{\langle X_2, X_1\rangle}{\langle X_1, X_1\rangle} X_1$을 빼 주면,

$$\big(X_2 - \tfrac{\langle X_2, X_1\rangle}{\langle X_1, X_1\rangle} X_1\big) \perp X_1$$

이 된다. 뿐만 아니라, 만약 $X_1 \perp X_2$였다면, X_3에서 X_2-성분 $\frac{\langle X_3, X_2\rangle}{\langle X_2, X_2\rangle} X_2$와 X_1-성분 $\frac{\langle X_3, X_1\rangle}{\langle X_1, X_1\rangle} X_1$을 빼 주면(그림도 그려 보라),

$$\big(X_3 - \tfrac{\langle X_3, X_2\rangle}{\langle X_2, X_2\rangle} X_2 - \tfrac{\langle X_3, X_1\rangle}{\langle X_1, X_1\rangle} X_1\big) \perp X_i, \qquad (i = 1, 2)$$

가 된다. 이를 일반화한 것이 Gram-Schmidt Orthogonalization이다.

정리 1.7 (Gram-Schmidt Orthogonalization) $\{X_1, \ldots, X_r\}$을 $\mathbb{R}^n$의 linearly independent subset이라고 하자.

(가) $Y_1 = X_1$으로, 그리고 $2 \le i \le r$일 때에는

$$Y_i = X_i - \tfrac{\langle X_i, Y_{i-1}\rangle}{\langle Y_{i-1}, Y_{i-1}\rangle} Y_{i-1} - \cdots - \tfrac{\langle X_i, Y_2\rangle}{\langle Y_2, Y_2\rangle} Y_2 - \tfrac{\langle X_i, Y_1\rangle}{\langle Y_1, Y_1\rangle} Y_1$$

으로 inductively 정의하면,

$$\langle X_1, \ldots, X_r\rangle = \langle Y_1, \ldots, Y_r\rangle$$

이고, $\{Y_1, \ldots, Y_r\}$은 $\mathbb{R}^n$의 orthogonal subset이 된다.

(나) 따라서, $\big\{\frac{1}{\|Y_1\|} Y_1, \ldots, \frac{1}{\|Y_r\|} Y_r\big\}$은 $\mathbb{R}^n$의 orthonormal subset이 된다.

(다) 특별히, $r = n$이면, $\big\{\frac{1}{\|Y_1\|} Y_1, \ldots, \frac{1}{\|Y_r\|} Y_r\big\}$은 $\mathbb{R}^n$의 orthonormal basis.

증명 : 귀납법. $\langle X_1, \ldots, X_i\rangle = \langle Y_1, \ldots, Y_i\rangle$이고, $\{Y_1, \ldots, Y_i\}$가 orthogonal subset인 것을 가정하고, $\langle X_1, \ldots, X_{i+1}\rangle = \langle Y_1, \ldots, Y_{i+1}\rangle$이고 $\{Y_1, \ldots, Y_{i+1}\}$도 orthogonal subset인 것을 보이면 된다. 독자들에게 맡긴다. ($Y_{i+1} \neq 0$임을 확인하는 것을 잊으면 안 된다.) □

다음 연습문제는 직관적으로 자명하지만, 분명히 확인해 둘 필요가 있다.

연습문제 1.8 위 정리 1.7에서 $\{X_1, \ldots, X_r\}$이 원래 orthogonal subset이면, 모든 i에 대해 $Y_i = X_i$임을 보여라.

다음 제 2 절에서 우리에게는 다음 따름정리가 필요하다.

따름정리 1.9 $\{X_1, \ldots, X_r\}$이 $\mathbb{R}^n$의 orthonormal subset이면, 이를 extend하여 $\mathbb{R}^n$의 orthonormal basis $\{X_1, \ldots, X_n\}$을 만들 수 있다.

증명 : 우선 $\mathfrak{C} = \{X_1, \ldots, X_r\}$이 orthonormal subset이므로, $\mathfrak{C}$는 일차독립이다(연습문제 1.6 (가)항). 이제 이를 확장해(Basis Extension Theorem) $\mathbb{R}^n$의 basis를 찾은 다음, Gram-Schmidt Orthogonalization을 이용하면 $\mathbb{R}^n$의 orthonormal basis $\{X_1, \ldots, X_n\}$을 찾을 수 있다(연습문제 1.8 참조). □

Euclidean space의 공부에는 반드시 (real) orthogonal group이 따라다닌다. (왜 그런지 알려면, "선형대수학 2"를 수강해야 한다.)

정의 1.10 Real **orthogonal group** $\mathbf{O}(n)$을

$$\mathbf{O}(n) = \{A \in \mathfrak{M}_{n,n}(\mathbb{R}) \mid A^{\mathbf{t}} \cdot A = I\}$$

로 정의한다. 우리는 $\mathbf{O}(n)$의 원소를 **orthogonal matrix**라고 부른다.

다음 연습문제는 [I, 연습문제 1.1.18]을 복사한 것이다.

연습문제 1.11 다음을 보여라.

(가) $I \in \mathbf{O}(n)$.
(나) $A, B \in \mathbf{O}(n)$이면, $AB \in \mathbf{O}(n)$.
(다) $A \in \mathbf{O}(n)$이면, (A는 가역이고) $A^{-1} \in \mathbf{O}(n)$.
(라) $A \in \mathbf{O}(n)$이면, $A^{\mathbf{t}} \in \mathbf{O}(n)$.

Orthogonal group의 공부는 끝이 없지만, 다음 제 2 절의 목적을 위해서는 다음 연습문제만으로 충분하다.

연습문제 1.12 다음 조건

(1) $A \in \mathbf{O}(n)$.
(2) A의 column들 $\{[A]^1, \ldots, [A]^n\}$은 $\mathbb{R}^n$의 orthonormal basis.

은 동치임을 보여라.

위 연습문제는 무엇보다도

$$[\text{orthogonal matrix를 찾는 문제}] = [\text{orthonormal basis를 찾는 문제}]$$

인 것을 말해 준다.

제 2 절 Real Symmetric Matrix 의 대각화

section version

150607

이 節에서는 real symmetric matrix 는 항상 diagonalizable 임을 증명한다. 우선 새로운 표기법과 terminology 부터 소개한다.

정의 2.1 $A=(a_{ij})\in\mathfrak{M}_{n,n}(\mathbb{C})$ 일 때,

$$\overline{A}=(\overline{a_{ij}}),\qquad A^*=\overline{A}^{\mathbf{t}}=\overline{A^{\mathbf{t}}}$$

로 정의하고, A^* 를 A 의 **adjoint matrix** 라고 부른다.[7]

이 節의 내용은 다음 관찰에 depend 하고 있다. 다음 관찰은 정말 simple 하고 또 거의 자명하지만, 의외로 심각한 뜻을 내포하고 있다. ("심각한 뜻"이란 ······, transpose matrix ($F=\mathbb{R}$ case) 와 adjoint matrix ($F=\mathbb{C}$ case) 의 '출생의 비밀' ·······.[8])

관찰 2.2 F^n 과 F^m 에 각각 dot product $\langle\ ,\rangle$ 가 주어져 있다고 하자.[9] 이때 모든 $A\in\mathfrak{M}_{m,n}(F)$ 에 대해,

$$\langle AX,Y\rangle=\big\langle X,A^*\cdot Y\big\rangle,\quad \langle Y,AX\rangle=\big\langle A^*\cdot Y,X\big\rangle,\qquad (X\in F^n,\ \ Y\in F^m)$$

이 성립한다. (물론, $F=\mathbb{R}$ 일 때에는 $A^*=A^{\mathbf{t}}$ 로 이해하면 된다.)

증명 : Dot product 의 정의로부터,

$$\langle AX,Y\rangle=(AX)^{\mathbf{t}}\cdot\overline{Y}=X^{\mathbf{t}}\cdot A^{\mathbf{t}}\cdot\overline{Y}$$

이고, 한편 $\overline{A^*}=A^{\mathbf{t}}$ 이므로,

$$\big\langle X,A^*\cdot Y\big\rangle=X^{\mathbf{t}}\cdot\overline{A^*\cdot Y}=X^{\mathbf{t}}\cdot\overline{A^*}\cdot\overline{Y}=X^{\mathbf{t}}\cdot A^{\mathbf{t}}\cdot\overline{Y}$$

이다. 두 번째 등식은 독자들에게 맡긴다. □

이제 real symmetric matrix 의 대각화를 시작해 보자. 그런데, $F=\mathbb{R}$ 인 경우에는 eigen-value 가 하나도 없을 수 있으므로, 대각화를 시작조차 할 수 없을지도 모른다. 따라서 우리는 무엇보다도 먼저 real symmetric matrix 는 (real) eigen-value 를 갖는 것부터 설명해야 한다.

[7] 이 adjoint 는 § 6.7 의 classical adjoint adj (A) 와는 무관.

[8] 이 comment 를 이해하려면, "선형대수학 2" 를 수강해야 한다.

[9] F 는 항상 $\mathbb{R}$ 또는 $\mathbb{C}$.

명제 2.3 Symmetric matrix $A \in \mathfrak{M}_{n,n}(\mathbb{R})$ 은 항상 (real) eigen-value 를 갖는다.

증명 : Real matrix A 를 complex matrix $A \in \mathfrak{M}_{n,n}(\mathbb{C})$ 로 이해한다고 해서 뭐 안 될 이유는 없다. 그리고 $\mathbb{C}^n$ 에는 Hermitian dot product 가 주어졌다고 생각하자. 이제, eigen-value $\lambda \in \mathbb{C}$ 를 갖는 A 의 eigen-vector $X \in \mathbb{C}^n$ 을 찾는다(즉, $AX = \lambda X$). 그러면, $A^* = A^{\mathbf{t}} = A$ 이므로, 관찰 2.2 에 의해

$$\lambda \langle X, X \rangle = \langle AX, X \rangle = \big\langle X, A^* \cdot X \big\rangle = \langle X, AX \rangle = \langle X, \lambda X \rangle = \overline{\lambda} \langle X, X \rangle$$

를 얻는다(마지막 등식은 연습문제 1.4 (나)항 참조). 따라서, $\overline{\lambda} = \lambda \in \mathbb{R}$. □

위 명제(의 증명)에 따르면, 모든 real symmetric matrix 의 characteristic polynomial 은 $\mathbb{R}$-위에서 1-차식들의 곱으로 분해됨을 알 수 있다. 왜냐하면, 그의 (complex) root 은 사실은 모두 real number 이기 때문이다.

준비가 하나 더 필요하다.

관찰 2.4 $\mathfrak{B}$ 가 $\mathbb{R}^n$ 의 orthonormal basis 이면, $[I]^{\mathfrak{B}}_{\mathcal{E}} \in \mathbf{O}(n)$.

증명 : 연습문제 1.12 에 의해 자명. □

이 article 의 main theorem 을 state 하기 전에 먼저 독자들의 '사고방식'을 점검할 필요가 있다.

관찰 2.5 $A \in \mathfrak{M}_{n,n}(\mathbb{R})$ 일 때, 다음 조건

(1) A 의 eigen-vector 들로 이루어진 $\mathbb{R}^n$ 의 **orthonormal** basis 존재.
(2) $U^{-1}AU$ 가 diagonal matrix 인 $U \in \mathbf{O}(n)$ 존재.

은 동치이다.

증명 : 만에 하나 "관찰 2.4 에 의해 자명하다"는 말 이외에 무언가 더 쓰고 싶은 독자가 있다면, [I, 관찰 7.2.13] 의 증명(들)을 (여러 번) 다시 읽어야만 할 것이다! □

우리가 real symmetric matrix 에 관심을 갖는 것은 다음 관찰 때문이다.

관찰 2.6 $A \in \mathfrak{M}_{n,n}(\mathbb{R})$ 일 때, 만약 A 의 eigen-vector 들로 이루어진 $\mathbb{R}^n$ 의 orthonormal basis 가 존재하면, A 는 symmetric.

증명 : 앞 관찰 2.5 에 의해, $U^{-1}AU = D$ 인 $U \in \mathbf{O}(n)$ 과 diagonal matrix D 가 존재할 것이다. 그런데, $U^{\mathbf{t}} = U^{-1}$ 이고 $D^{\mathbf{t}} = D$ 이므로,

$$D = D^{\mathbf{t}} = (U^{-1}AU)^{\mathbf{t}} = U^{-1}A^{\mathbf{t}} \cdot U$$

가 된다. 따라서 $U^{-1}AU = U^{-1}A^{\mathbf{t}} \cdot U$ 이므로, $A = A^{\mathbf{t}}$. (이 증명은 한마디로 "대각행렬은 대칭행렬".) □

그리고, 위 관찰의 converse 도 성립한다. 우리는 다음 정리를 Spectral Theorem (의 하나) 이라고 부른다.[10] 독자들은 지금 [I, 정리 7.3.1] 의 증명 technique 을 복습하기 바란다.

정리 2.7 (Spectral Theorem : Real Symmetric Matrix Case) $A \in \mathfrak{M}_{n,n}(\mathbb{R})$ 일 때, 다음 조건은 동치이다.

(1) A 는 symmetric.

(2) A 의 eigen-vector 들로 이루어진 $\mathbb{R}^n$ 의 orthonormal basis 존재.

증명 : (i) 한쪽 방향은 이미 증명하였으므로, 이제 A 는 symmetric 이라고 가정하고, **matrix size n 에 관한 귀납법**을 사용해 $U^{-1}AU$ 가 diagonal matrix 인 $U \in \mathbf{O}(n)$ 이 존재함을 보이자. 우선 $n = 1$ 이면 증명할 것이 없으므로, $n \geq 2$ 라고 가정하자.

(ii) 당연히, 먼저 A 의 eigen-vector $X_1 \in \mathbb{R}^n$ 을 하나 찾는다 (명제 2.3). 이때 우리는 X_1 이 unit vector 라고 생각할 수 있다 (왜 그런가?). 그리고, X_1 의 eigen-value 를 $\lambda \in \mathbb{R}$ 이라고 놓자 (즉, $AX_1 = \lambda X_1$). 다음에는, Gram-Schmidt Orthogonalization 을 이용해 $\mathbb{R}^n$ 의 orthonormal basis $\mathfrak{B} = \{X_1, X_2, \ldots, X_n\}$ 을 찾는다 (따름정리 1.8 참조). 이제 $U_0 = [I]^{\mathfrak{B}}_{\mathcal{E}} \in \mathbf{O}(n)$ 으로 표기하면 (관찰 2.4 참조), L_A 의 행렬 표현 $[L_A]^{\mathfrak{B}}_{\mathfrak{B}}$ 는

$$[L_A]^{\mathfrak{B}}_{\mathfrak{B}} = U_0^{-1}AU_0 = \left(\begin{array}{c|c} \lambda & C \\ \hline 0 & B \end{array}\right)$$

의 형태가 된다 (단, $B \in \mathfrak{M}_{n-1,n-1}(\mathbb{R})$, $C \in \mathfrak{M}_{1,n-1}(\mathbb{R})$).

---

[10] 전통적으로, square matrix (또는 linear operator) 의 [eigen-value 들의 집합] 을 spectrum 이라고 부른다. 그렇지만, 'Spectral Theorem' 의 뜻은 분명치 않다. 그저 그렇게 命名된 정리들을 통틀어 'Spectral Theorems' 라고 부른다고 이해하면 된다. 혹은, diagonalizable operator 의 'spectrum' 이 diagonal matrix 의 대각선에 나타난다고 생각해도 그럴듯할 (?) 것이다. (Spectrum 에 관한 이 해석에는 별로 自信이 없다.)

(iii) 그런데, ${U_0}^{\mathbf{t}} = U_0^{-1}$ 이고 $A^{\mathbf{t}} = A$ 이므로,

$$\left(\begin{array}{c|c} \lambda & 0 \\ \hline C^{\mathbf{t}} & B^{\mathbf{t}} \end{array}\right) = \left(\begin{array}{c|c} \lambda & C \\ \hline 0 & B \end{array}\right)^{\mathbf{t}} = (U_0^{-1}AU_0)^{\mathbf{t}} = U_0^{-1}AU_0 = \left(\begin{array}{c|c} \lambda & C \\ \hline 0 & B \end{array}\right)$$

가 된다. 즉, B 는 **symmetric** 이고 $C = 0$ 이다. 이제 induction hypothesis 에 의해, $P^{-1}BP$ 가 diagonal matrix 인 $P \in \mathbf{O}(n-1)$ 이 존재할 것이다. 마지막으로 $U_1 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P \end{array}\right)$ 로 놓으면, 당연히 $U_1 \in \mathbf{O}(n)$ 이고 $U_1^{-1} = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P^{-1} \end{array}\right)$ 이므로,

$$U_1^{-1}U_0^{-1}AU_0U_1 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P^{-1} \end{array}\right)\left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & B \end{array}\right)\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P \end{array}\right) = \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & P^{-1}BP \end{array}\right)$$

가 된다. 그런데, $P^{-1}BP$ 는 diagonal matrix 이므로, $U = U_0U_1 \in \mathbf{O}(n)$ 으로 놓으면(연습문제 1.11(나)항 참조), $U^{-1}AU$ 는 diagonal matrix. □

따라서, real symmetric matrix 는 단순히 대각화가 가능할 뿐만 아니라, 이때 orthogonal transition matrix 를 찾을 수 있다. 이 상황은 orthonormal basis 를 '직교좌표계'로 번역하면 의미가 더 분명해진다.[11] 즉, real symmetric matrix 는 적당한 '직교좌표계'에 대해 각각의 좌표축에 'diagonally act (operate)'하는 것들이다.

조심해야 할 사항이 하나 있다. 즉, complex symmetric matrix 는 대각화가 불가능할 수도 있다. (정의 2.1 다음 단락에서 잠시 암시가 있었지만, $F = \mathbb{C}$ 일 때, real symmetric matrix 에 대응하는 개념은 self-adjoint matrix 이다.[12])

연습문제 2.8 연습문제 7.2.11(나)항을 다시 풀어 보라. 즉, $\begin{pmatrix} 2 & \mathbf{i} \\ \mathbf{i} & 4 \end{pmatrix}$ 는 대각화가 불가능함을 보여라.

제 3 절 Application

section version 180724

Real symmetric matrix 의 Spectral Theorem (정리 2.7) 에는 수많은 'classical' application 들이 있다. 그중 가장 중요한 것은 2-차형식의 대각화이다.

[11] 엄밀하게는 '정규직교좌표계'.
[12] $A^* = A$ 인 $A \in \mathfrak{M}_{n,n}(\mathbb{C})$ 를 self-adjoint matrix 라고 부른다.

이제, n-개의 실변수 $x_1, \dots, x_n$ 을 갖는 '同次인 2-차식 (quadratic form)'

$$2\,Q\big((x_1, \dots, x_n)^{\mathbf{t}}\big) = \sum_i a_{ii} x_i^2 + 2 \sum_{i<j} a_{ij} x_i x_j$$

가 주어졌다고 하자 (단, $a_{ij} \in \mathbb{R}$). 이때, $J = (a_{ij})$ 로 놓으면 (단, $a_{ij} = a_{ji}$),

$$2\,Q(X) = X^{\mathbf{t}} \cdot JX, \qquad (X \in \mathbb{R}^n)$$

임을 쉽게 확인할 수 있다. 그런데, J 는 symmetric matrix 이므로, Spectral Theorem 에 의해

$$U^{-1} J\, U = D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$$

인 $U \in \mathbf{O}(n)$ 이 존재할 것이다 (단, $\lambda_i \in \mathbb{R}$). 따라서, $U^{\mathbf{t}} = U^{-1}$ 이므로,

$$2\,Q(X) = X^{\mathbf{t}}(UDU^{-1})X = (U^{-1}X)^{\mathbf{t}} \cdot D(U^{-1}X)$$

가 된다. 이제, $U^{-1}X = X' = (x_1', \dots, x_n')^{\mathbf{t}}$ 로 표기하면 (즉, $X = UX'$),

$$2\,Q(X) = (X')^{\mathbf{t}} \cdot DX' = \lambda_1 (x_1')^2 + \cdots + \lambda_n (x_n')^2$$

의 '예쁜' 모양을 얻을 수 있다. 이때, 우리는 quadratic form $2\,Q$ 를 **대각화**했다고 말한다. (이 논의의 point 는 $U^{\mathbf{t}} = U^{-1}$.)

보기 3.1 Symmetric matrix $J = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ 에 대응하는 quadratic form

$$2\,Q(x, y) = 2x^2 + 2xy + 2y^2$$

이 주어졌다고 하자. 이제, J 를 대각화하면,

$$\tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \cdot \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

이 된다. 따라서,

$$\tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

즉,

$$x = \tfrac{1}{\sqrt{2}}\, x' + \tfrac{1}{\sqrt{2}}\, y', \qquad y = -\tfrac{1}{\sqrt{2}}\, x' + \tfrac{1}{\sqrt{2}}\, y'$$

으로 치환하면,

$$2x^2 + 2xy + 2y^2 = (x')^2 + 3(y')^2$$

을 얻는다.

보기 3.2 이번에는 2-차곡선

$$2x^2+2xy+2y^2=4$$

의 graph를 그려 보자. 이제, 앞 보기에서 J의 eigen-vector들로 이루어진 $\mathbb{R}^2$의 (ordered) orthonormal basis

$$\mathfrak{B}=\{v,w\}=\left\{\tfrac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix},\ \tfrac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}\right\}$$

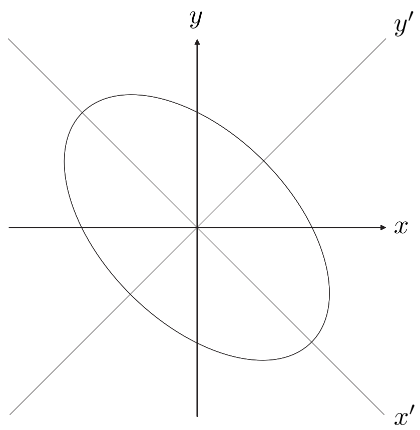
을 생각하면, $U=[I]^{\mathfrak{B}}_{\mathcal{E}}$ 이므로,

$$[X]_{\mathcal{E}}=X=UX'=[I]^{\mathfrak{B}}_{\mathcal{E}}\cdot[x'v+y'w]_{\mathfrak{B}}=[x'v+y'w]_{\mathcal{E}}$$

가 된다(앞 보기의 표기법 참조). 따라서, $X=x'v+y'w$ 이므로, [표준기저 $\mathcal{E}$에 관한 좌표가 $(x,y)^{\mathbf{t}}$ 인 점]은 기저 $\mathfrak{B}$에 관한 좌표가 $(x',y')^{\mathbf{t}}$ 가 된다. 그런데, 주어진 곡선은 '(x',y')-좌표'로

$$(x')^2+3(y')^2=4$$

이고, 따라서 곡선의 graph는 다음



과 같이 '보통 타원'을 시계 방향으로 45°-회전시킨 모양이 된다.

연습문제 3.3 모든 실수 x,y에 대하여,

$$x^2+xy+y^2-x-y+\frac{1}{3}\geq 0$$

임을 증명하라.[13]

[13]물론, 판별식만으로도 충분하지만…….

연습문제 3.4 다음 2-차곡선의 graph 를 그려라.

(가) $2x^2 - 2\sqrt{2}\,xy + 3y^2 = 4$.

(나) $x^2 - 6xy + y^2 = 2$.

(다) $11x^2 + 24xy + y^2 = 15$.

연습문제 3.5 우리가 고등학교에서 배운 2-차곡선은 타원(원 포함)과 쌍곡선 그리고 포물선뿐이다. 왜 그것들만 배웠을까?[14]

위에서 우리는 linear operator/matrix 가 아닌 quadratic form 의 대각화(즉, J 의 대각화)를 논하였다. 사실 [I] 에서 저자는 이 부분을 애써 숨기고 있다고 할 수 있다. 그러나, 지금은 $U \in \mathbf{O}(n)$ 이면, $U^{-1} = U^{\mathbf{t}}$ 인 특수한 경우이므로 문제의 본질을 피할 수 있다.[15]

Quadratic form 의 대각화와 관련된 보기를 하나만 더 생각해 보자. 다음 연습문제는 [I, 연습문제 15.3.11] 을 elementary language 로 번역한 것이다. 독자들은 Spectral Theorem 없이는 다음 연습문제에 답하는 것이 쉽지 않음을 음미하기 바란다.

연습문제 3.6 $J \in \mathfrak{M}_{n,n}(\mathbb{R})$ 이 symmetric 일 때(단, $\langle\ ,\ \rangle$ 는 $\mathbb{R}^n$ 의 dot product),

(가) 다음 조건

(1) $0 \neq X \in \mathbb{R}^n$ 이면, $\langle JX, X\rangle > 0$.

(2) J 의 eigen-value 는 모두 positive real number.

은 동치임을 증명하라. 이때 J 를 **positive definite matrix** 라고 부른다.

(나) 다음 조건

(1) $0 \neq X \in \mathbb{R}^n$ 이면, $\langle JX, X\rangle \geq 0$.

(2) J 의 eigen-value 는 모두 non-negative real number.

은 동치임을 증명하라. 이때 J 를 **positive semi-definite matrix** 라고 부른다.

(다) J 가 positive semi-definite 이면, $K^2 = J$ 인 $K \in \mathfrak{M}_{n,n}(\mathbb{R})$ 이 존재함을 보여라.

[14] 이 연습문제의 '완벽'한 답을 위하여 너무 많은 시간을 허비하지 말 것.

[15] 본질적인 논의는 몇 년 후로 미룬다.