3장 부록

보충설명: 3.1절

> 정리 3A.1. 사영 행렬과 부분공간 분해

$n \times n$ 행렬 P에 대해 다음이 성립한다.

(i) 사영 행렬의 성질

P가 **사영(projection)** 행렬이면, $I - P$도 사영 행렬이 되며 다음 관계가 성립한다.

$$\mathcal{R}(I-P) = \mathcal{N}(P), \quad \mathcal{R}(P) = \mathcal{N}(I-P) \tag{1}$$

즉, P의 행 공간(range)과 $I - P$의 영공간(null space)이 같으며, $I - P$의 행 공간과 P의 영공간이 같다.

(ii) 부분공간의 직교 분해

벡터 공간 $\mathbb{R}^n$의 두 부분공간 $S_1 = \mathcal{R}(P)$, $S_2 = \mathcal{N}(P)$에 대해,

- $S_1 \cap S_2 = \{0\}$
- $S_1 \oplus S_2 = \mathbb{R}^n$

즉, 임의의 벡터 $\boldsymbol{v} \in \mathbb{R}^n$는 $\boldsymbol{v}_1 = P\boldsymbol{v} \in S_1$, $\boldsymbol{v}_2 = (I-P)\boldsymbol{v} \in S_2$ 를 사용하여

$$\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2 \tag{2}$$

로 **고유하게 분해**할 수 있다.

(증명)

(i) $I - P$가 사영 행렬임을 보임: 먼저,

$$(I-P)^2 = I - 2P + P^2 = I - P \tag{3}$$

이므로 $I - P$도 사영 행렬이다. 다음으로,

- $P\boldsymbol{x} \in \mathcal{R}(P)$,
- $(I-P)\boldsymbol{x} \in \mathcal{R}(I-P)$

이고,

$$P(P\boldsymbol{x}) = P\boldsymbol{x} \quad \Rightarrow \quad P[(I-P)\boldsymbol{x}] = (I-P)[P\boldsymbol{x}] = 0 \tag{4}$$

이므로,

- $P\boldsymbol{x} \in \mathcal{N}(I-P)$
- $(I-P)\boldsymbol{x} \in \mathcal{N}(P)$

따라서,

$$\mathcal{R}(I-P) \subset \mathcal{N}(P), \quad \mathcal{R}(P) \subset \mathcal{N}(I-P) \tag{5}$$

또한, 만약 $\boldsymbol{x} \in \mathcal{N}(P)$, 즉 $P\boldsymbol{x} = 0$이면, $(I-P)\boldsymbol{x} = \boldsymbol{x} \in \mathcal{R}(I-P)$ 이므로

$$\mathcal{N}(P) \subset \mathcal{R}(I-P) \tag{6}$$

마찬가지로 $\mathcal{N}(I-P) \subset \mathcal{R}(P)$도 성립한다.

(ii) 벡터 공간의 직교 분해: 벡터 $\boldsymbol{v}$는 다음과 같이 분해된다.

$$\boldsymbol{v} = P\boldsymbol{v} + (I-P)\boldsymbol{v} \tag{7}$$

여기서, $P\boldsymbol{v} \in \mathcal{R}(P) = S_1$이고, (i)에 의해 $(I-P)\boldsymbol{v} \in \mathcal{R}(I-P) = \mathcal{N}(P) = S_2$이다.

또한, $\boldsymbol{v} \in \mathcal{N}(I-P) \cap \mathcal{N}(P)$이면,

$$\boldsymbol{v} = \boldsymbol{v} - P\boldsymbol{v} = (I-P)\boldsymbol{v} = 0 \tag{8}$$

이므로,

$$S_1 \cap S_2 = \mathcal{N}(I-P) \cap \mathcal{N}(P) = \{0\} \tag{9}$$

유일성을 보이기 위해, 만일 $\boldsymbol{v}_1 = P\boldsymbol{v} + \boldsymbol{v}_3 \in S_1$이라 하면, $\boldsymbol{v}_1 + \boldsymbol{v}_2 = \boldsymbol{v}$ 이므로

$$\boldsymbol{v}_2 = (I-P)\boldsymbol{v} - \boldsymbol{v}_3 \in S_2 \tag{10}$$

가 되고, $\boldsymbol{v}_3 \in S_1 \cap S_2 = \{0\}$ 이므로 유일하게 분해됨을 보일 수 있다. ¶

보조정리 3A.2. 정사영 행렬과 직교 성질

$n \times n$ 행렬 P에 대해 다음이 성립한다.

(i) 정사영 행렬의 필요충분조건

P가 **정사영(orthogonal projection)** 행렬이 될 필요충분조건은 다음과 같다.

$$\mathcal{R}(P) \perp \mathcal{N}(P) \tag{11}$$

즉, P의 행 공간(range)과 영공간(null space)이 서로 직교한다.

(ii) $I - P$도 정사영 행렬

P가 정사영 행렬이면, $I-P$도 정사영 행렬이 되며 다음 관계가 성립한다.

$$\mathcal{R}(I-P) \perp \mathcal{R}(P) \tag{12}$$

즉, $I-P$의 행 공간과 P의 행 공간이 서로 직교한다.

(증명)

(i) 정사영 행렬의 필요충분조건:

($\Rightarrow$) P가 정사영 행렬이면, $P=P^T$이므로 선형대수의 기본정리에 의해 다음이 성립한다.

$$\mathcal{R}(P) \perp \mathcal{N}(P^T) \Rightarrow \mathcal{R}(P) \perp \mathcal{N}(P) \tag{13}$$

($\Leftarrow$) SVD를 이용한 증명: $S_1=\mathcal{R}(P)$의 정규직교 기저를 $\{\boldsymbol{q}_1,\ldots,\boldsymbol{q}_k\}$, $S_2=\mathcal{N}(P)$의 정규직교 기저를 $\{\boldsymbol{q}_{k+1},\ldots,\boldsymbol{q}_n\}$라고 하면,

- $P\boldsymbol{q}_i=\boldsymbol{q}_i$, $i=1,\ldots,k$
- $P\boldsymbol{q}_j=0$, $j=k+1,\ldots,n$

따라서,

$$PQ = P\left[\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n\right] = \left[\boldsymbol{q}_1,\ldots,\boldsymbol{q}_k,\mathbf{0},\ldots,\mathbf{0}\right] \tag{14}$$

$$Q^TPQ = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} \quad \Rightarrow \quad P = Q\begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}Q^T \tag{15}$$

이므로 $P^T=P$가 성립한다.

(ii) $I-P$도 정사영 행렬임을 보임:

$$(I-P)^T = I-P \tag{16}$$

이므로 정리 3-1.1에 의해 쉽게 유도된다. ¶

보충 설명: 3.2.2절

하우스홀더 QR 분해 (Householder QR factorization)

1. 정사영 행렬과 하우스홀더 반사 행렬

주어진 벡터 $\boldsymbol{v}$에 대해, 부분공간 $S = \mathrm{span}(\boldsymbol{v})$ 위로의 정사영(orthogonal projection) 행렬과 그 직교 여공간 $S^{\perp}$ 위로의 정사영 행렬은 각각 다음과 같이 정의된다.

$$P_v = \frac{\boldsymbol{v}\boldsymbol{v}^T}{\boldsymbol{v}^T\boldsymbol{v}}, \quad P_v^{\perp} = I - P_v = I - \frac{\boldsymbol{v}\boldsymbol{v}^T}{\boldsymbol{v}^T\boldsymbol{v}} \tag{17}$$

이제, **하우스홀더 반사(Householder reflector)** 행렬은 다음과 같이 정의된다.

$$H_v = I - 2P_v = I - 2\frac{\boldsymbol{v}\boldsymbol{v}^T}{\boldsymbol{v}^T\boldsymbol{v}} \tag{18}$$

이때, $P_v^{\perp}$는 랭크 $(m-1)$을 가지는 반면, 하우스홀더 반사 행렬 H_v는 **꽉 찬 랭크(full rank)를 가지는 직교 행렬**이다. 즉, 다음 성질을 만족한다.

$$H_v^T H_v = I \tag{19}$$

하우스홀더 반사는 주어진 벡터를 특정 방향으로 반사(reflection)하는 변환을 수행하며, 이를 이용해 행렬을 상삼각 행렬로 변환할 수 있다.

2. 하우스홀더 반사를 이용한 벡터 변환

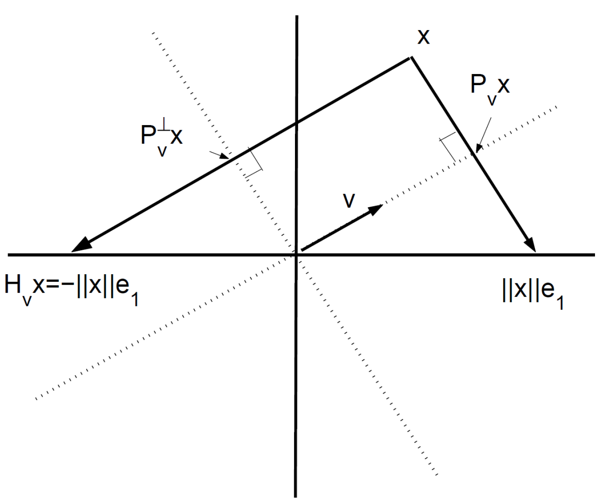

그림 3A.1: 하우스홀더 반사를 이용한 벡터 변환

위 그림에서 보이는 **하우스홀더 반사(Householder reflector)**는 주어진 벡터의 특정 요소를 제외한 나머지 요소들을 0으로 변환하는 데 사용된다. 즉, 주어진 벡터 $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T$에 대해, 다음과 같이 특정 방향으로의 반사를 통해 원하는 형태의 변환을 수행할 수 있다.

$$\boldsymbol{x} \longrightarrow H_v\boldsymbol{x} = \alpha\boldsymbol{e}_1 \tag{20}$$

이때, H_v는 직교 행렬(orthogonal matrix)이므로, 벡터의 2-노름이 보존된다.

$$\|\boldsymbol{x}\|_2 = \|H_v\boldsymbol{x}\|_2 \tag{21}$$

따라서, $\alpha = \pm\|\boldsymbol{x}\|_2$이고, 이를 통해 $\boldsymbol{v}$를 찾을 수 있다.

$$H_v\boldsymbol{x} = \boldsymbol{x} - 2\frac{\boldsymbol{v}^T\boldsymbol{x}}{\boldsymbol{v}^T\boldsymbol{v}}\boldsymbol{v} = \pm\|\boldsymbol{x}\|_2\boldsymbol{e}_1 \tag{22}$$

이를 만족하는 $\boldsymbol{v}$는 다음과 같이 선택할 수 있다.

$$\boldsymbol{v} = \boldsymbol{x} \mp \|\boldsymbol{x}\|_2\boldsymbol{e}_1 \tag{23}$$

하지만, 만약 $\boldsymbol{x}$가 $\boldsymbol{e}_1$ 방향에 가깝다면, $\boldsymbol{v} = \boldsymbol{x} - \mathrm{sign}(x_1)\|\boldsymbol{x}\|_2\boldsymbol{e}_1$은 작은 노름을 가지므로 상대적인 오류가 커질 수 있다. 이를 방지하기 위해, 다음과 같이 $\boldsymbol{v}$를 선택하는 것이 일반적이다.

$$\boldsymbol{v} = \boldsymbol{x} + \mathrm{sign}(x_1)\|\boldsymbol{x}\|_2\boldsymbol{e}_1 \tag{24}$$

3. 하우스홀더 QR 분해 과정

QR 분해에서는 LU 분해와 유사하게 **순차적인 직교 변환을 통해 행렬 A를 상삼각 행렬로 변환**한다. 예를 들어, A가 4×3 행렬이라고 하자.

(1) 첫 번째 단계

$$A = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \quad \Longrightarrow \quad A^{(1)} = Q_1 A = \begin{bmatrix} \alpha_1 & * & * \\ 0 & * & * \\ 0 & * & * \\ 0 & * & * \end{bmatrix} \tag{25}$$

이때, 하우스홀더 행렬 Q_1은 다음과 같다.

$$Q_1 = H_{v_1}, \quad \boldsymbol{v}_1 = A(:,1) + \mathrm{sign}(A_{11})\|A(:,1)\|_2\boldsymbol{e}_1 \tag{26}$$

(2) 두 번째 단계

$$A^{(1)} \quad \Longrightarrow \quad A^{(2)} = Q_2 A^{(1)} = \begin{bmatrix} \alpha_1 & * & * \\ 0 & \alpha_2 & * \\ 0 & 0 & * \\ 0 & 0 & * \end{bmatrix} \tag{27}$$

이때, Q_2는 다음과 같이 정의된다.

$$Q_2 = \begin{bmatrix} 1 & \mathbf{0} \\ 0 & H_{v_2} \end{bmatrix} = I - 2\frac{\overline{\boldsymbol{v}}_2 \overline{\boldsymbol{v}}_2^T}{\overline{\boldsymbol{v}}_2^T \overline{\boldsymbol{v}}_2} \tag{28}$$

여기서,

$$\overline{\boldsymbol{v}}_2 = \begin{bmatrix} 0 \\ \boldsymbol{v}_2 \end{bmatrix}, \quad \boldsymbol{v}_2 = A^{(1)}(2:m,2) + \text{sign}(A^{(1)}_{22}) \|A^{(1)}(2:m,2)\|_2 \boldsymbol{e}_2(2:m) \tag{29}$$

(3) 세 번째 단계

$$A^{(2)} \quad \Longrightarrow \quad A^{(3)} = Q_3 A^{(2)} = \begin{bmatrix} \alpha_1 & * & * \\ 0 & \alpha_2 & * \\ 0 & 0 & \alpha_3 \\ 0 & 0 & 0 \end{bmatrix} \tag{30}$$

이때, Q_3는 다음과 같이 정의된다.

$$Q_3 = \begin{bmatrix} 1 & 0 & \mathbf{0} \\ 0 & 1 & 0 \\ 0 & 0 & H_3 \end{bmatrix} = I - 2\frac{\overline{\boldsymbol{v}}_3 \overline{\boldsymbol{v}}_3^T}{\overline{\boldsymbol{v}}_3^T \overline{\boldsymbol{v}}_3} \tag{31}$$

여기서,

$$\overline{\boldsymbol{v}}_3 = \begin{bmatrix} 0 \\ 0 \\ \boldsymbol{v}_3 \end{bmatrix}, \quad \boldsymbol{v}_3 = A^{(2)}(3:m,3) + \text{sign}(A^{(2)}_{33}) \|A^{(2)}(3:m,3)\|_2 \boldsymbol{e}_3(3:m) \tag{32}$$

즉, $Q_3 Q_2 Q_1 A = R$이며, 직교 행렬 Q는 $Q = Q_1^T Q_2^T Q_3^T$를 만족한다. 따라서, $A = QR$ **형태의 QR 분해를 얻을 수 있다**.

일반적인 $m \times n$ 행렬에 대한 하우스홀더 QR 분해

일반적인 $m \times n$ 행렬 A에 대해, **하우스홀더 QR 알고리즘**을 수행하려면 n개의 직교 행렬 Q_j $(j = 1, 2, \ldots, n)$을 생성해야 한다. 이때, 각 Q_j는 다음과 같은 구조를 가진다.

$$Q_j = \begin{bmatrix} I & 0 \\ 0 & H_j \end{bmatrix} \tag{33}$$

여기서, H_j는 **하우스홀더 반사 행렬(Householder reflector)**이며, 다음과 같이 정의된다.

$$H_j = I - 2\frac{\overline{\boldsymbol{v}}_j \overline{\boldsymbol{v}}_j^T}{\overline{\boldsymbol{v}}_j^T \overline{\boldsymbol{v}}_j} \tag{34}$$

하우스홀더 반사 행렬을 정의하기 위해, 벡터 $\overline{\boldsymbol{v}}_j$를 다음과 같이 설정한다.

$$\overline{\boldsymbol{v}}_j = \begin{bmatrix} 0 \\ \boldsymbol{v}_j \end{bmatrix} \tag{35}$$

여기서, $\boldsymbol{v}_j$는 행렬 A의 j번째 열을 기반으로 계산되며, 이 벡터를 사용하여 하우스홀더 변환을 수행하게 된다.

하우스홀더 QR 알고리즘의 과정

하우스홀더 QR 분해는 **순차적인 직교 변환을 통해 행렬 A를 상삼각 행렬 R로 변환**하는 과정이다.

1. 초기 행렬 A에서 첫 번째 열을 변환
    - 첫 번째 열을 **하우스홀더 반사**를 통해 첫 번째 단위 벡터 방향으로 정렬
    - 첫 번째 직교 행렬 Q_1 생성
2. 두 번째 열을 변환
    - Q_1을 적용한 후, 두 번째 열을 하우스홀더 변환하여 정렬
    - 두 번째 직교 행렬 Q_2 생성
3. 이 과정을 반복하여 상삼각 행렬 R을 생성
    - Q_n까지 순차적으로 수행

이 과정을 통해 최종적으로 $A = QR$ 형태의 분해를 얻을 수 있다.

$$Q_n \dots Q_2 Q_1 A = R \tag{36}$$

직교 행렬 Q는 개별 행렬 Q_j들의 전치 행렬 곱으로 주어진다.

$$Q = Q_1^T Q_2^T \dots Q_n^T \tag{37}$$

기븐스 회전을 사용한 QR 분해 (QR factorization using Givens rotations)

2차원 회전 행렬 (Rotation Matrix)

주어진 벡터 $\boldsymbol{x} = (x_1, x_2)^T$에 대해, 2차원 **회전 행렬(rotation matrix)** G_θ는 다음과 같이 정의된다.

$$G_\theta = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}, \quad c = \cos\theta, \quad s = \sin\theta, \quad \theta = \tan^{-1}\frac{x_2}{x_1} \tag{38}$$

이 행렬은 **직교 행렬(orthogonal matrix)**이며, 적절한 θ 값을 선택하면 하우스홀더 반사와 유사하게 다음 변환을 수행할 수 있다.

$$G_\theta \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ 0 \end{bmatrix} \tag{39}$$

즉, 기븐스 회전은 특정 벡터의 한 성분을 **0으로 변환하는 과정**을 제공한다.

일반적인 기븐스 회전 (Givens Rotation in Higher Dimensions)

일반적인 m차원 공간에서, 특정한 두 개의 좌표 (i, k)에 대해 **기븐스 회전(Givens rotation)** 행렬 $G_\theta(i, k)$는 다음과 같이 정의된다.

$$G_\theta(i,k) = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & c & \dots & s & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & -s & \dots & c & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}, \quad c = \cos\theta, \quad s = \sin\theta \tag{40}$$

이때, $G_\theta(i, k)$는 **직교 행렬**이다. 기븐스 회전은 특정 열에 대해 행렬 연산을 수행할 때, 한 번의 연산으로 하나의 요소를 0으로 만들 수 있다. 따라서, 기븐스 회전은 **희소 행렬(sparse matrix)**의 연산에 적합하며, 수치적으로 매우 안정적인 방법이다.

기븐스 회전을 이용한 QR 분해 과정

아래는 4×3 행렬 A에 대한 기븐스 회전 과정을 나타낸다.

$$\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} \xrightarrow{G(4,1)} \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \mathbf{0} & \times & \times \end{bmatrix} \xrightarrow{G(3,1)} \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \mathbf{0} & \times & \times \\ 0 & \times & \times \end{bmatrix} \tag{41}$$

$$\xrightarrow{G(2,1)} \begin{bmatrix} \times & \times & \times \\ \mathbf{0} & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix} \xrightarrow{G(4,2)} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \\ 0 & \mathbf{0} & \times \end{bmatrix} \tag{42}$$

$$\xrightarrow{G(3,2)} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \mathbf{0} & \times \\ 0 & 0 & \times \end{bmatrix} \xrightarrow{G(4,3)} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \\ 0 & 0 & \mathbf{0} \end{bmatrix} \tag{43}$$

기븐스 QR 분해의 수행 과정

이 4×3 행렬에서, QR 분해 과정은 다음과 같은 순서로 이루어진다.

$$G(4,3)G(3,2)G(4,2)G(2,1)G(3,1)G(4,1)A = R \tag{44}$$

그리고, 직교 행렬 Q는 다음과 같이 구할 수 있다.

$$Q = G(4,1)^T G(3,1)^T G(2,1)^T G(4,2)^T G(3,2)^T G(4,3)^T \tag{45}$$

즉, 기븐스 회전은 **하나씩 특정 요소를 제거하면서 직교 행렬을 순차적으로 적용하는 방식**을 따른다.

결론:

- 기븐스 회전은 **특정 열의 요소를 0으로 만들기 위해 순차적으로 적용하는 직교 변환 기법**이다.
- 기븐스 회전을 사용한 QR 분해는 **희소 행렬 연산에서 매우 효율적**이며, **수치적으로 안정적인 방식**이다.
- 기븐스 회전은 하우스홀더 QR 분해와 비교하여, **국소적인 연산을 수행하는 특성**이 있으며, **병렬 연산에서 더 적합한 경우가 많다**.

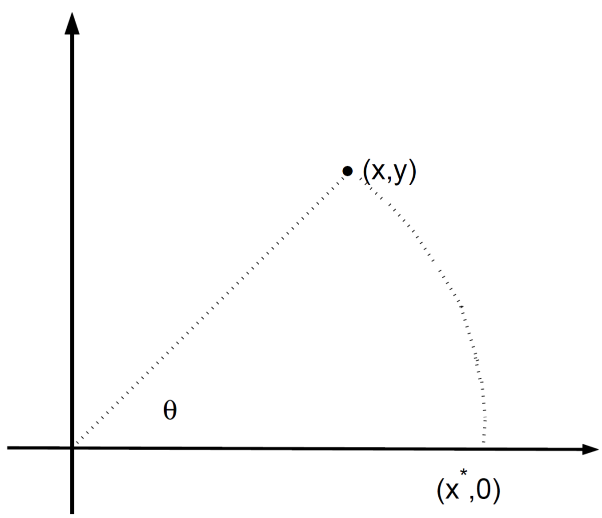

그림 3A.2: 벡터의 기븐스 회전

각주