# 행렬대수 1 - 사영과 QR 분해

시립대학교 통계학과

2019년 6월 3일

## 1 두 벡터의 사영

선형독립인 두 벡터 $\boldsymbol{a}_1$과 $\boldsymbol{a}_2$ 가 있다고 하자. 벡터 $\boldsymbol{a}_1$과 같은 방향을 가지면서 벡터 $\boldsymbol{a}_2$에 가장 가까운 벡터를 $proj_{\boldsymbol{a}_1}(\boldsymbol{a}_2)$ 라고 정의하고 이를 벡터 $\boldsymbol{a}_1$ 방향으로 벡터 $\boldsymbol{a}_2$ 의 사영(projection) 이라고 부른다.

그러면 이러한 사영은 어떻게 구할 수 있나? 벡터 $\boldsymbol{a}_2$ 의 사영은 벡터 $\boldsymbol{a}_1$ 방향에 있으므로 어떤 실수 $c$ 가 있어서 다음과 같이 표시할 수 있다.

$$proj_{\boldsymbol{a}_1}(\boldsymbol{a}_2) = c\boldsymbol{a}_1$$

이제 사영 $c\boldsymbol{a}_1$과 벡터 $\boldsymbol{a}_2$의 거리 $d(c)$ 를 생각하면 다음과 같다.

$$\begin{aligned} d^2(c) &= \|\boldsymbol{a}_2 - c\boldsymbol{a}1\|^2 \\ &= (\boldsymbol{a}_2 - c\boldsymbol{a}_1)^t(\boldsymbol{a}_2 - c\boldsymbol{a}_1) \\ &= \boldsymbol{a}_2^t\boldsymbol{a}_2^t - 2c\boldsymbol{a}_2^t\boldsymbol{a}_1 + c^2\boldsymbol{a}_1^t\boldsymbol{a}_1^t \end{aligned}$$

위에서 $\|\boldsymbol{a}\|$ 는 벡터 $\boldsymbol{a}$의 길이를 나타낸다.

$$d(\boldsymbol{a}) = \|\boldsymbol{a}\| = \sqrt{\boldsymbol{a}^t\boldsymbol{a}}$$

상수 $c$ 는 거리 $d(c)$를 최소로 만드는 수이다. $d^2(c)$은 $c$ 에 대하여 미분 가능한 2차 함수이며 아래로 볼록한 함수이므로 이를 미분하여 $c$ 를 구할 수 있다.

$$\frac{\partial d^2(c)}{\partial c} = -2\boldsymbol{a}_2^t\boldsymbol{a}_1 + 2c\boldsymbol{a}_1^t\boldsymbol{a}_1^t = 0$$

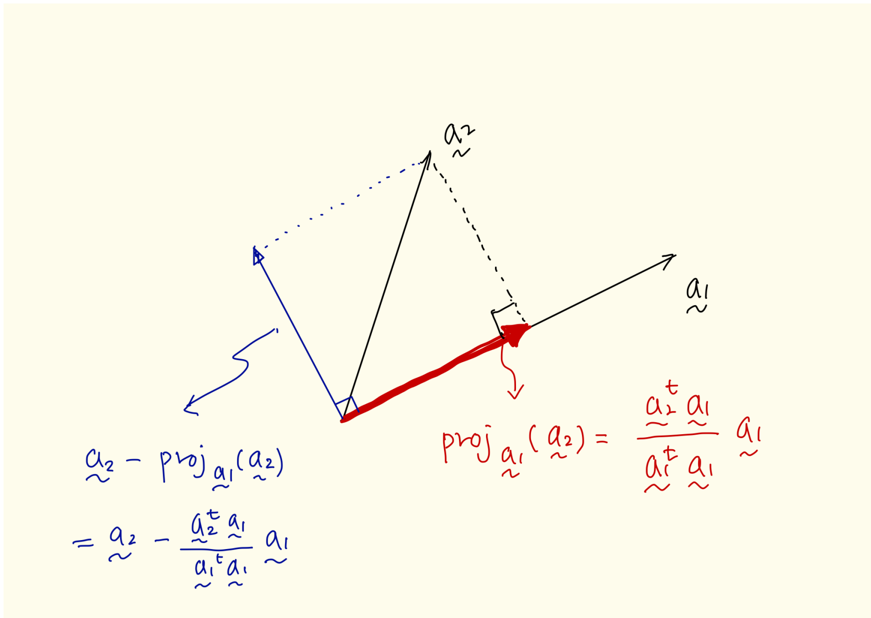


Figure 1: "벡터의 사영

위의 방적식으로 부터 $c$를 얻고

$$c = \frac{a_2^t a_1}{a_1^t a_1^t}$$

다음과 같이 벡터 $a_1$ 방향으로 벡터 $a_2$ 의 사영을 나타낼 수 있다.

$$proj_{a_1}(a_2) = \frac{a_2^t a_1}{a_1^t a_1} a_1 \tag{1}$$

이제 위의 두 벡터의 사영을 이용하면 벡터 $a_1$ 과 직교하는 벡터 $\tilde{q}_2$를 다음과 같이 찾을 수 있다.

$$\tilde{q}_2 = a_2 - proj_{a_1}(a_2) = a_2 - \frac{a_2^t a_1}{a_1^t a_1} a_1$$

두 벡터 $a_1$와 $\tilde{q}_2$의 직교성은 다음과 같이 보일 수 있다.

$$
\begin{aligned}
\boldsymbol{a}_1^t \tilde{\boldsymbol{q}}_2 &= \boldsymbol{a}_1^t \left( \boldsymbol{a}_2 - \frac{\boldsymbol{a}_2^t \boldsymbol{a}_1}{\boldsymbol{a}_1^t \boldsymbol{a}_1} \boldsymbol{a}_1 \right) \\
&= \boldsymbol{a}_1^t \boldsymbol{a}_2 - \frac{\boldsymbol{a}_2^t \boldsymbol{a}_1}{\boldsymbol{a}_1^t \boldsymbol{a}_1} \boldsymbol{a}_1^t \boldsymbol{a}_1 \\
&= \boldsymbol{a}_1^t \boldsymbol{a}_2 - \boldsymbol{a}_2^t \boldsymbol{a}_1 \\
&= 0
\end{aligned} \tag{2}
$$

이제 두 벡터 $\boldsymbol{q}_1$과 $\boldsymbol{q}_2$ 를 다음과 같이 정규직교벡터로 만들 수 있다.

$$
\begin{aligned}
\boldsymbol{q}_1 &= \boldsymbol{a}_1 / \|\boldsymbol{a}_1\| \\
\boldsymbol{q}_2 &= \tilde{\boldsymbol{q}}_2 / \|\tilde{\boldsymbol{q}}_2\|
\end{aligned}
$$

## 1.1 Gram–Schmidt 방법

서로 독립인 $n$차원의 벡터들이 $p$개 있을때

$$
\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_p
$$

이들이 만드는 열공간을 $C$라고 하자.

$$
\begin{aligned}
C &= span\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_p\} \\
&= \{\ c_1 \boldsymbol{a}_1 + c_2 \boldsymbol{a}_2 + \cdots + c_p \boldsymbol{a}_p \mid \text{ all possible real values of } c_1, c_2, \ldots, c_p\ \}
\end{aligned} \tag{3}
$$

이제 우리는 위와 동일한 열공간 $C$ 만드는 정규직교 벡터들을 찾는 방법을 알아보고자 한다.

$$
\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_p \quad \text{where } \boldsymbol{q}_i^t \boldsymbol{q}_j = 0, \ \ \boldsymbol{q}_i^t \boldsymbol{q}_i = 1
$$

그리고

$$
S = span\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_p\} = span\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_p\} \tag{4}
$$

이제 앞 절의 벡터의 사영에 대한 결과를 사용하여 다음과 같은 직교하는 $p$ 개의 벡터들을 축차적으로 만들어 보자.

$$
\begin{aligned}
\tilde{\boldsymbol{q}}_1 &= \boldsymbol{a}_1 \\
\tilde{\boldsymbol{q}}_2 &= \boldsymbol{a}_2 - proj_{\tilde{\boldsymbol{q}}_1}(\boldsymbol{a}_2) \\
\tilde{\boldsymbol{q}}_3 &= \boldsymbol{a}_3 - proj_{\tilde{\boldsymbol{q}}_1}(\boldsymbol{a}_3) - proj_{\tilde{\boldsymbol{q}}_2}(\boldsymbol{a}_3) \\
\tilde{\boldsymbol{q}}_4 &= \boldsymbol{a}_4 - proj_{\tilde{\boldsymbol{q}}_1}(\boldsymbol{a}_4) - proj_{\tilde{\boldsymbol{q}}_2}(\boldsymbol{a}_4) - proj_{\tilde{\boldsymbol{q}}_3}(\boldsymbol{a}_4) \\
&\cdots \\
\tilde{\boldsymbol{q}}_p &= \boldsymbol{a}_p - \sum_{k=1}^{p} proj_{\tilde{\boldsymbol{q}}_k}(\boldsymbol{a}_p) \qquad (5)
\end{aligned}
$$

축차적으로 만든 벡터들을 정규벡터로 만들면 원래의 벡터들 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_p$이 생성하는 동일한 열공간을 만드는 정규직교 벡터 $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_p$를 만들 수 있다.

$$
\boldsymbol{q}_i = \tilde{\boldsymbol{q}}_i / \|\tilde{\boldsymbol{q}}_i\|, \quad i = 1, 2, \ldots, p \qquad (6)
$$

Gram–Schmidt 방법으로 만든 벡터들이 직교하는 것은 다음과 같이 증명할 수 있다. 먼저 식 (2) 에 의하여 $\tilde{\boldsymbol{q}}_1$ 과 $\tilde{\boldsymbol{q}}_2$는 직교한다. 이제 임의의 $i$에 대하여 $\tilde{\boldsymbol{q}}_1, \tilde{\boldsymbol{q}}_2, \cdots, \tilde{\boldsymbol{q}}_{i-1}$ 벡터들이 직교한다고 가정하자. 모든 $1 \leq j \leq i-1$ 에 대하여

$$
\begin{aligned}
\tilde{\boldsymbol{q}}_j^t \tilde{\boldsymbol{q}}_i &= \tilde{\boldsymbol{q}}_j^t \left[ \boldsymbol{a}_i - \sum_{k=1}^{i} proj_{\tilde{\boldsymbol{q}}_k}(\boldsymbol{a}_i) \right] \\
&= \tilde{\boldsymbol{q}}_j^t \left[ \boldsymbol{a}_i - proj_{\tilde{\boldsymbol{q}}_j}(\boldsymbol{a}_i) \right] - \left[ \sum_{\substack{1 \leq k \leq i-1 \\ k \neq j}} \tilde{\boldsymbol{q}}_j^t \, proj_{\tilde{\boldsymbol{q}}_k}(\boldsymbol{a}_i) \right] \\
&= 0 + 0
\end{aligned}
$$

위에서 마지막 단계의 직교성은 다음과 같은 사실로 부터 유도된다.

1. $\boldsymbol{a}_i - proj_{\tilde{\boldsymbol{q}}_j}(\boldsymbol{a}_i)$는 $\tilde{\boldsymbol{q}}_j^t$와 직교한다.

2. 가정에 의하여 $\tilde{\boldsymbol{q}}_1, \tilde{\boldsymbol{q}}_2, \cdots, \tilde{\boldsymbol{q}}_{i-1}$는 직교하고 $proj_{\tilde{\boldsymbol{q}}_k}(\boldsymbol{a}_i)$는 $\tilde{\boldsymbol{q}}_k$ 와 같은 방향을 가진다.

$$
\tilde{\boldsymbol{q}}_j^t proj_{\tilde{\boldsymbol{q}}_k}(\boldsymbol{a}_i) = 0 \quad \text{for } 1 \leq j, k \leq i-1, k \neq j
$$

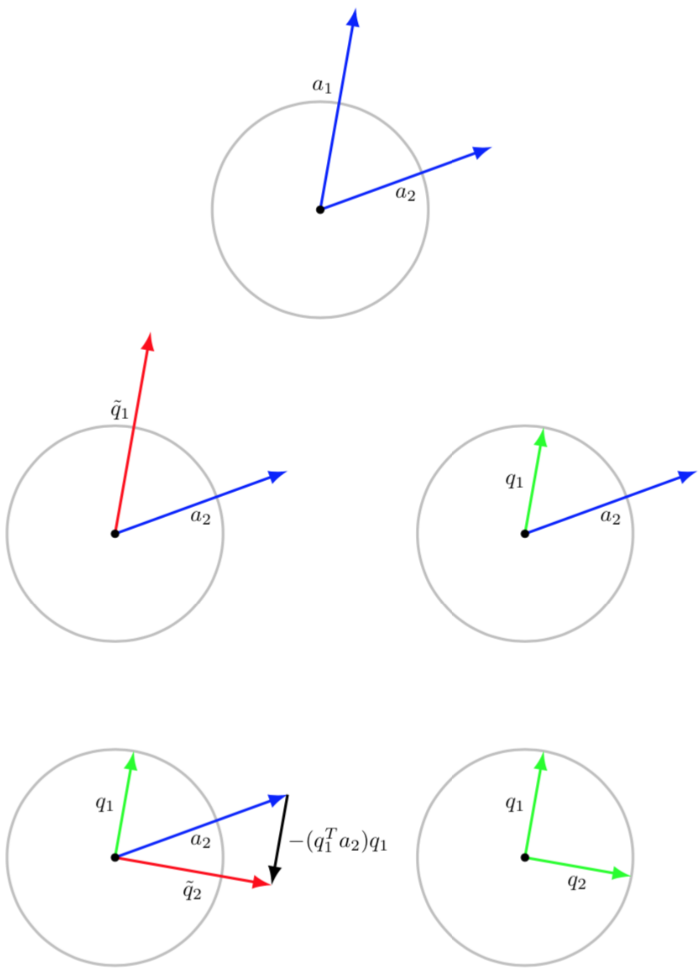


**Figure 5.3** Gram–Schmidt algorithm applied to two 2-vectors $a_1$, $a_2$. *Top.* The original vectors $a_1$ and $a_2$. The gray circle shows the points with norm one. *Middle left.* The orthogonalization step in the first iteration yields $\tilde{q}_1 = a_1$. *Middle right.* The normalization step in the first iteration scales $\tilde{q}_1$ to have norm one, which yields $q_1$. *Bottom left.* The orthogonalization step in the second iteration subtracts a multiple of $q_1$ to yield the vector $\tilde{q}_2$, which is orthogonal to $q_1$. *Bottom right.* The normalization step in the second iteration scales $\tilde{q}_2$ to have norm one, which yields $q_2$.

Figure 2: Gram–Schmidt 방법을 설명한 그림(출처:Introduction to Applied Linear Algebra by Boyd and Vandenberghe, 2019)

식 (5)과 (6)의 알고리즘을 Gram–Schmidt 방법이라고 부른다. 위의 두 식에 의한 알고리즘을 다음과 같은 사실을 이용하면 좀 더 간단한 방법의 알고리즘이 나온다.

$$proj_{\tilde{\boldsymbol{q}}_k}(\boldsymbol{a}_l) = \frac{\boldsymbol{a}_l^t \tilde{\boldsymbol{q}}_k}{\tilde{\boldsymbol{q}}_k^t \tilde{\boldsymbol{q}}_k} \tilde{\boldsymbol{q}}_k = \frac{\boldsymbol{a}_l^t \tilde{\boldsymbol{q}}_k}{\|\tilde{\boldsymbol{q}}_k\|^2} \tilde{\boldsymbol{q}}_k = (\boldsymbol{a}_l^t \boldsymbol{q}_k)\boldsymbol{q}_k$$

- $p$개의 벡터 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_p$에 대하여
- $for\ i = 1, 2, \ldots, p$

1. $\tilde{\boldsymbol{q}}_i = \boldsymbol{a}_i - (\boldsymbol{q}_1^t \boldsymbol{a}_i)\boldsymbol{q}_1 - \cdots - (\boldsymbol{q}_{i-1}^t \boldsymbol{a}_i)\boldsymbol{q}_{i-1}$ (직교화)
2. $\boldsymbol{q}_i = \tilde{\boldsymbol{q}}_i / \|\boldsymbol{q}_i\|$ (정규화)

## 2 최소제곱법과 사영

회귀계수벡터의 값을 구하는 최소제곱법의 기준을 다시 살펴보자.

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} (y_i - \boldsymbol{x}_i^t \boldsymbol{\beta})^2 = \min_{\boldsymbol{\beta}} (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^t (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) = \min_{\boldsymbol{\beta}} \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 \tag{7}$$

위에서 $\boldsymbol{X}\boldsymbol{\beta}$는 행렬 $\boldsymbol{X}$의 열벡터 $\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_p$ 로 이루어진 선형조합이다.

$$\boldsymbol{X}\boldsymbol{\beta} = [\boldsymbol{a}_0 \ \boldsymbol{a}_1 \ \cdots \boldsymbol{a}_p]\boldsymbol{\beta} = \beta_0 \boldsymbol{a}_0 + \beta_1 \boldsymbol{a}_1 + \cdots + \beta_p \boldsymbol{a}_p$$

따라서 최소제곱법으로 구한 회귀계수 벡터 $\hat{\boldsymbol{\beta}}$는 반응값 벡터 $\boldsymbol{y}$ 와 $\boldsymbol{X}\boldsymbol{\beta}$의 거리가 최소가 되도록 만들어 준다.

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^t \boldsymbol{X})^{-1} \boldsymbol{X}^t \boldsymbol{y}, \quad \min_{\boldsymbol{\beta}} \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 = \left\|\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\right\|^2$$

따라서 예측값 벡터 $\hat{\boldsymbol{y}}$ 는 행렬 $\boldsymbol{X}$의 열벡터로 생성한 열공간 방향으로 반응값 벡터 $\boldsymbol{y}$를 사영한 벡터이다.

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}(\boldsymbol{X}^t \boldsymbol{X})^{-1} \boldsymbol{X}^t \boldsymbol{y}$$

위에서 행렬 $\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}^t \boldsymbol{X})^{-1} \boldsymbol{X}^t$를 열공간 $C(\boldsymbol{X})$의 사영행렬(projection matirx)라고 부른다.

## 3 QR 분해

식 (5)과 (6)에 주어진 Gram–Schmidt 방법을 원래 벡터들 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_p$에 대하여 다시 다음과 같이 나타낼 수 있다.

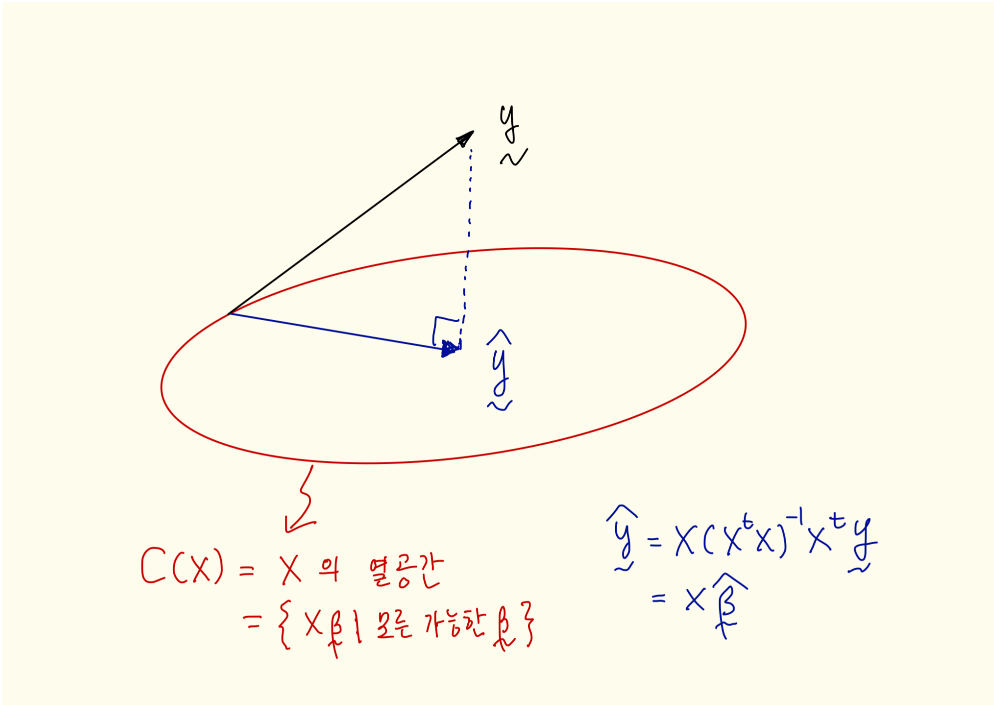


Figure 3: 최소제곱법을 설명한 그림

$$
\begin{aligned}
\boldsymbol{a}_1 &= \tilde{\boldsymbol{q}}_1 \\
&= \|\tilde{\boldsymbol{q}}_1\| \, \boldsymbol{q}_1 \\
\boldsymbol{a}_2 &= \tilde{\boldsymbol{q}}_2 + proj_{\tilde{\boldsymbol{q}}_1}(\boldsymbol{a}_2) \\
&= \tilde{\boldsymbol{q}}_2 + \frac{\boldsymbol{a}_2^t \tilde{\boldsymbol{q}}_1}{\tilde{\boldsymbol{q}}_1^t \tilde{\boldsymbol{q}}_1} \tilde{\boldsymbol{q}}_1 \\
&= (\boldsymbol{a}_2^t \boldsymbol{q}_1)\boldsymbol{q}_1 + \|\tilde{\boldsymbol{q}}_2\| \, \boldsymbol{q}_2 \\
\boldsymbol{a}_3 &= \tilde{\boldsymbol{q}}_3 + proj_{\tilde{\boldsymbol{q}}_1}(\boldsymbol{a}_3) + proj_{\tilde{\boldsymbol{q}}_2}(\boldsymbol{a}_3) \\
&= \tilde{\boldsymbol{q}}_3 + \frac{\boldsymbol{a}_3^t \tilde{\boldsymbol{q}}_1}{\tilde{\boldsymbol{q}}_1^t \tilde{\boldsymbol{q}}_1} \tilde{\boldsymbol{q}}_1 + \frac{\boldsymbol{a}_3^t \tilde{\boldsymbol{q}}_2}{\tilde{\boldsymbol{q}}_2^t \tilde{\boldsymbol{q}}_2} \tilde{\boldsymbol{q}}_2 \\
&= (\boldsymbol{a}_3^t \boldsymbol{q}_1)\boldsymbol{q}_1 + (\boldsymbol{a}_3^t \boldsymbol{q}_2)\boldsymbol{q}_2 + \|\tilde{\boldsymbol{q}}_3\| \, \boldsymbol{q}_3 \\
&\dots \\
\boldsymbol{a}_p &= (\boldsymbol{a}_p^t \boldsymbol{q}_1)\boldsymbol{q}_1 + (\boldsymbol{a}_p^t \boldsymbol{q}_2)\boldsymbol{q}_2 + \cdots + (\boldsymbol{a}_p^t \boldsymbol{q}_{p-1})\boldsymbol{q}_{p-1} + \|\tilde{\boldsymbol{q}}_p\| \, \boldsymbol{q}_p
\end{aligned}
$$

즉 위의 축차식을 보면 원래 벡터 $\boldsymbol{a}_i$ 는 Gram–Schmidt 방법으로 구한 정규직교벡터 $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_p$ 의 선형 조합으로 나타낼 수 있다.

이제 Gram–Schmidt 방법으로 구한 정규직교벡터들 $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_p$ 을 모아놓은 행렬을 $\boldsymbol{Q}$ 라고 하고 위에서 $\boldsymbol{a}_i$들이 직교행렬의 선형조합으로 표시될때 계수들을 모아놓는 상삼각행렬을 $\boldsymbol{R}$ 이라고 하자. 그러면 다음과 같은 QR 분해가 주어진다.

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R} \tag{8}$$

여기서

$$\boldsymbol{Q} = [\boldsymbol{q}_1 \;\; \boldsymbol{q}_2 \;\; \ldots \;\; \boldsymbol{q}_p], \quad \boldsymbol{Q}^t\boldsymbol{Q} = \boldsymbol{I}$$
$$\boldsymbol{R} = \begin{bmatrix} \|\tilde{\boldsymbol{q}}_1\| & \boldsymbol{a}_2^t\boldsymbol{q}_1 & \boldsymbol{a}_3^t\boldsymbol{q}_1 & \ldots & \boldsymbol{a}_p^t\boldsymbol{q}_1 \\ 0 & \|\tilde{\boldsymbol{q}}_2\| & \boldsymbol{a}_3^t\boldsymbol{q}_2 & \ldots & \boldsymbol{a}_p^t\boldsymbol{q}_2 \\ 0 & 0 & \|\tilde{\boldsymbol{q}}_3\| & \ldots & \boldsymbol{a}_p^t\boldsymbol{q}_3 \\ & & & \ldots & \\ 0 & 0 & 0 & \ldots & \|\tilde{\boldsymbol{q}}_p\| \end{bmatrix}$$

# 4 최소제곱법과 QR 분해

다시 식 (7)의 최소제곱법을 고려하자. 실제로 $\hat{\boldsymbol{\beta}}$을 구하는 경우 많은 통계계산 프로그램에서 실제로 역행렬 $(\boldsymbol{X}^t\boldsymbol{X})^{-1}$을 계산하는 경우는 드물다. 실제 최소제곱법을 푸는 알고리즘으로 많은 경우 QR 분해법을 사용한다.

$n \times (p+1)$ 행렬 $\boldsymbol{X}$의 QR 분해가 다음과 같이 주어졌다고 하자. 또한 $r(\boldsymbol{X}) = p+1 < n$ 이라고 가정하자.

$$\boldsymbol{X} = \boldsymbol{Q}_1\boldsymbol{R}_1$$

그러면 행렬 $\boldsymbol{X}$는 다음과 같은 확장된 분해를 가진다.

$$\boldsymbol{X} = \boldsymbol{Q}\boldsymbol{R} = \begin{bmatrix} \boldsymbol{Q}_1 & \boldsymbol{Q}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{R}_1 \\ \boldsymbol{0} \end{bmatrix}$$

위에서 $\boldsymbol{Q}_2$ 는 행렬 $\boldsymbol{Q}_1$의 열들과 직교하는 $n-(p+1)$ 개의 추가 정규직교 벡터들로 이루어진 행렬이다. $\boldsymbol{Q}_1$과 $\boldsymbol{Q}_2$는 각각 $n \times (p+1), n \times (n-p-1)$의 행렬이다. 따라서 행렬 $\boldsymbol{Q} = [\boldsymbol{Q}_1 \; \boldsymbol{Q}_2]$ 는 $n \times n$ 직교행렬이다 ($\boldsymbol{Q}^t\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}^t = \boldsymbol{I}$). 여기서 $\boldsymbol{R}_1$는 $(p+1) \times (p+1)$ 상삼각행렬(upper diagonal matrix)고 $\mathbf{0}$은 차원이 $(n-p-1) \times (p+1)$인 영행렬이다.

이제 $\boldsymbol{Q}^t\boldsymbol{Q} = \boldsymbol{I}$를 이용하여 잔차제곱합 $\|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2$를 다음과 같이 분해할 수 있다.

$$
\begin{aligned}
\|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 &= (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^t(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) \\
&= (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^t\boldsymbol{Q}\boldsymbol{Q}^t(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) \\
&= (\boldsymbol{Q}^t\boldsymbol{y} - \boldsymbol{Q}^t\boldsymbol{X}\boldsymbol{\beta})^t(\boldsymbol{Q}^t\boldsymbol{y} - \boldsymbol{Q}^t\boldsymbol{X}\boldsymbol{\beta}) \\
&= (\boldsymbol{c} - \boldsymbol{R}\boldsymbol{\beta})^t(\boldsymbol{c} - \boldsymbol{R}\boldsymbol{\beta}) \\
&= \left( \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \end{bmatrix} - \begin{bmatrix} \boldsymbol{R}_1 \\ \mathbf{0} \end{bmatrix} \boldsymbol{\beta} \right)^t \left( \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \end{bmatrix} - \begin{bmatrix} \boldsymbol{R}_1 \\ \mathbf{0} \end{bmatrix} \boldsymbol{\beta} \right) \\
&= (\boldsymbol{c}_1 - \boldsymbol{R}_1\boldsymbol{\beta})^t(\boldsymbol{c}_1 - \boldsymbol{R}_1\boldsymbol{\beta}) + \boldsymbol{c}_2^t\boldsymbol{c}_2 \\
&= \|\boldsymbol{c}_1 - \boldsymbol{R}_1\boldsymbol{\beta}\|^2 + \|\boldsymbol{c}_2\|^2
\end{aligned} \tag{9}
$$

여기서

$$
\boldsymbol{c} = \boldsymbol{Q}^t\boldsymbol{y} = \begin{bmatrix} \boldsymbol{Q}_1^t \\ \boldsymbol{Q}_2^t \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} \boldsymbol{Q}_1^t\boldsymbol{y} \\ \boldsymbol{Q}_2^t\boldsymbol{y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \end{bmatrix}
$$

위의 식 (9)를 보면 벡터 $\boldsymbol{c}_2 = \boldsymbol{Q}_2^t\boldsymbol{y}$는 $\boldsymbol{\beta}$와 관계가 없으므로 잔차제곱합 $\|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2$ 을 최소화하는 $\boldsymbol{\beta}$는 $\|\boldsymbol{c}_1 - \boldsymbol{R}_1\boldsymbol{\beta}\|^2$을 0으로 만드는 것이다. 즉 $\boldsymbol{R}_1\boldsymbol{\beta} = \boldsymbol{c}_1$를 만족하는 $\boldsymbol{\beta}$가 최소제곱 추정량이다.

$$
\min_{\boldsymbol{\beta}} \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 = \min_{\boldsymbol{\beta}} \|\boldsymbol{c}_1 - \boldsymbol{R}_1\boldsymbol{\beta}\|^2 + \|\boldsymbol{c}_2\|^2
$$

$\boldsymbol{X}$가 완전계수 행렬이므로 상삼각행렬인 $\boldsymbol{R}_1$도 완전계수이다. 따라서 방정식 $\boldsymbol{R}_1\boldsymbol{\beta} = \boldsymbol{c}_1$는 유일한 해는 상삼각행렬의 성질을 이용하여 축차식으로 쉽게 구할 수 있다.

$$
\hat{\boldsymbol{\beta}} = \boldsymbol{R}_1^{-1}\boldsymbol{c}_1
$$

위의 추정량은 실제 역행렬을 구하지 않고 QR 분해를 이용하여 회귀계수를 구하는 방법 중에 하나이다.

# 5 예제

## 5.1 Gram–Schmidt 방법

아래와 같이 4차원 벡터 3개가 있다.

$$\boldsymbol{a}_1 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \quad \boldsymbol{a}_2 = \begin{bmatrix} -1 \\ 3 \\ -1 \\ 3 \end{bmatrix} \quad \boldsymbol{a}_3 = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix} \tag{10}$$

위의 벡터 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$에 대하여 Gram–Schmidt 방법을 적용해보자.

- $i = 1$. 먼저 $\|\tilde{\boldsymbol{q}}_1\| = \|\boldsymbol{a}_1\| = 2$이므로 첫번째 벡터 $\boldsymbol{q}_1$를 만든다.

$$\boldsymbol{q}_1 = \tilde{\boldsymbol{q}}_1 / \|\tilde{\boldsymbol{q}}_1\| = \begin{bmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix}$$

- $i = 2$. 이제 두번째 직교벡터 $\boldsymbol{q}_2$를 만들자. $\boldsymbol{q}_1^t \boldsymbol{a}_2 = 4$이므로

$$\tilde{\boldsymbol{q}}_2 = \boldsymbol{a}_2 - (\boldsymbol{q}_1^t \boldsymbol{a}_2)\boldsymbol{q}_1 = \begin{bmatrix} -1 \\ 3 \\ -1 \\ 3 \end{bmatrix} - 4 \begin{bmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

  그리고 $\|\tilde{\boldsymbol{q}}_2\| = 2$이므로

$$\boldsymbol{q}_2 = \tilde{\boldsymbol{q}}_2 / \|\tilde{\boldsymbol{q}}_2\| = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$$

- $i = 3$ 마지막으로 $\boldsymbol{q}_1^t \boldsymbol{a}_3 = 2$, $\boldsymbol{q}_2^t \boldsymbol{a}_3 = 8$ 이므로

$$\tilde{\boldsymbol{q}}_3 = \boldsymbol{a}_3 - (\boldsymbol{q}_1^t \boldsymbol{a}_3)\boldsymbol{q}_1 - (\boldsymbol{q}_2^t \boldsymbol{a}_3)\boldsymbol{q}_2 = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix} - 2 \begin{bmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix} - 8 \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \\ 2 \\ 2 \end{bmatrix}$$

또한 $\|\tilde{\boldsymbol{q}}_3\| = 4$이므로

$$\boldsymbol{q}_3 = \tilde{\boldsymbol{q}}_3 / \|\tilde{\boldsymbol{q}}_3\| = \begin{bmatrix} -1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$$

따라서 Gram–Schmidt 방법으로 만든 정규직교벡터는 다음과 같다.

$$\boldsymbol{q}_1 = \begin{bmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix} \quad \boldsymbol{q}_2 = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} \quad \boldsymbol{q}_3 = \begin{bmatrix} -1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$$

# 6 QR 분해

식 (10) 에서 주어진 백터들을 열로 가진 행렬 $\boldsymbol{A}$의 QR분해를 구해보자.

$$\boldsymbol{A} = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix}$$

앞의 예제에서 구한 직교벡터를 그대로 이용하면 $\boldsymbol{Q}$는 쉽게 구해진다.

$$\boldsymbol{Q} = \begin{bmatrix} -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 \end{bmatrix}$$

또한 식 (8)에 주어진 공식을 이용하면 행렬 $\boldsymbol{R}$은 다음과 같이 구할 수 있다.

$$\boldsymbol{R} = \begin{bmatrix} \|\tilde{\boldsymbol{q}}_1\| & \boldsymbol{a}_2^t\boldsymbol{q}_1 & \boldsymbol{a}_3^t\boldsymbol{q}_1 \\ 0 & \|\tilde{\boldsymbol{q}}_2\| & \boldsymbol{a}_3^t\boldsymbol{q}_2 \\ 0 & 0 & \|\tilde{\boldsymbol{q}}_3\| \end{bmatrix} = \begin{bmatrix} 2 & 4 & 2 \\ 0 & 2 & 8 \\ 0 & 0 & 4 \end{bmatrix}$$