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.

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

위의 방적식으로 부터 c를 얻고

$$c = \frac{\boldsymbol{a}_2^t \boldsymbol{a}_1}{\boldsymbol{a}_1^t \boldsymbol{a}_1^t}$$

다음과 같이 벡터 $\boldsymbol{a}_1$ 방향으로 벡터 $\boldsymbol{a}_2$ 의 사영을 나타낼 수 있다.

$$proj_{\boldsymbol{a}_1}(\boldsymbol{a}_2) = \frac{\boldsymbol{a}_2^t \boldsymbol{a}_1}{\boldsymbol{a}_1^t \boldsymbol{a}_1} \boldsymbol{a}_1 \tag{1}$$

이제 위의 두 벡터의 사영을 이용하면 벡터 $\boldsymbol{a}_1$ 과 직교하는 벡터 $\tilde{\boldsymbol{q}}_2$를 다음과 같이 찾을 수 있다.

$$\tilde{\boldsymbol{q}}_2 = \boldsymbol{a}_2 - proj_{\boldsymbol{a}_1}(\boldsymbol{a}_2) = \boldsymbol{a}_2 - \frac{\boldsymbol{a}_2^t \boldsymbol{a}_1}{\boldsymbol{a}_1^t \boldsymbol{a}_1} \boldsymbol{a}_1$$

두 벡터 $\boldsymbol{a}_1$와 $\tilde{\boldsymbol{q}}_2$의 직교성은 다음과 같이 보일 수 있다.

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

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

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

1.1 Gram–Schmidt 방법

서로 독립인 n차원의 벡터들이 p개 있을때

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

이들이 만드는 열공간을 C라고 하자.

$$S = \{\ c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_p\boldsymbol{a}_p \mid \text{ all possible real values of } c_1, c_2, \ldots, c_p\ \} = span\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_p\} \tag{2}$$

이제 우리는 위와 동일한 열공간 S 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

$$S = span\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_p\} = span\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_p\} \tag{3}$$

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

$$\begin{aligned}
\tilde{\boldsymbol{q}}_1 &= \boldsymbol{a}_1 \\
\tilde{\boldsymbol{q}}_2 &= \boldsymbol{a}_2 - proj_{\tilde{\boldsymbol{q}}_1}(\boldsymbol{a}_2) \\
\tilde{\boldsymbol{q}}_3 &= \boldsymbol{a}_3 - proj_{\tilde{\boldsymbol{q}}_1}(\boldsymbol{a}_3) - proj_{\tilde{\boldsymbol{q}}_2}(\boldsymbol{a}_3) \\
\tilde{\boldsymbol{q}}_4 &= \boldsymbol{a}_4 - proj_{\tilde{\boldsymbol{q}}_1}(\boldsymbol{a}_4) - proj_{\tilde{\boldsymbol{q}}_2}(\boldsymbol{a}_4) - proj_{\tilde{\boldsymbol{q}}_3}(\boldsymbol{a}_4) \\
&\cdots \\
\tilde{\boldsymbol{q}}_p &= \boldsymbol{a}_p - \sum_{k=1}^{p} proj_{\tilde{\boldsymbol{q}}_k}(\boldsymbol{a}_p)
\end{aligned} \tag{4}$$

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.

$$\boldsymbol{q}_i = \tilde{\boldsymbol{q}}_i / \|\tilde{\boldsymbol{q}}_i\|, \quad i = 1, 2, \ldots, p \tag{5}$$

식 (4)과 (5)의 알고리즘을 Gram–Schmidt 방법이라고 부른다.

2 최소제곱법과 사영

회귀계수벡터의 값을 구하는 최소제곱법의 기준을 다시 살펴보자.

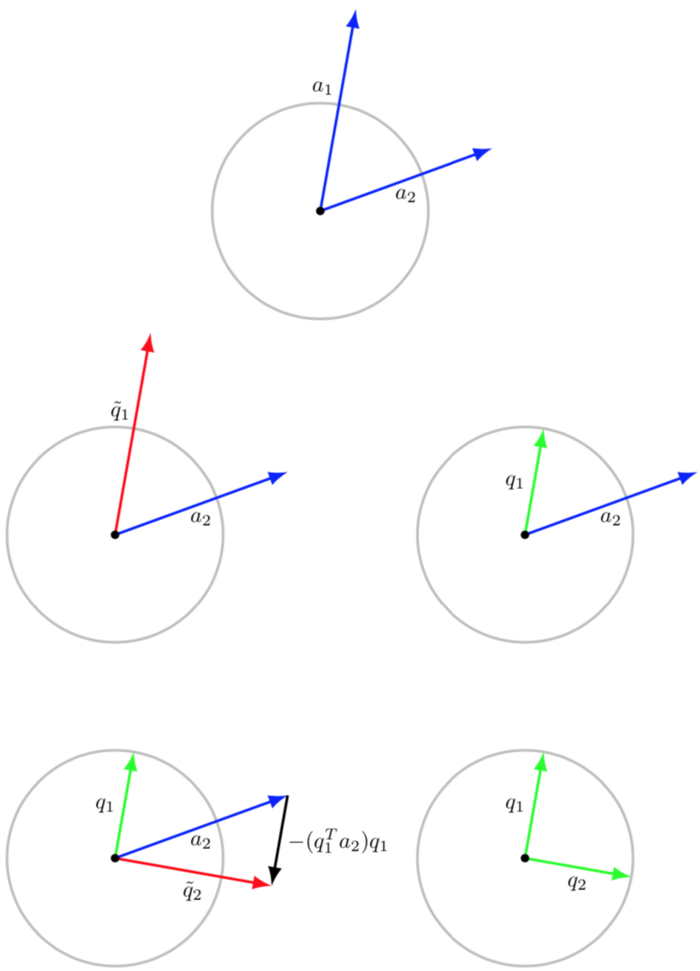

Figure 5.3 Gram–Schmidt algorithm applied to two 2-vectors a_1, a_2. *Top.* The original vectors a_1 and a_2. The gray circle shows the points with norm one. *Middle left.* The orthogonalization step in the first iteration yields $\tilde{q}_1 = a_1$. *Middle right.* The normalization step in the first 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 1: "Gram–Schmidt 방법을 설명한 그림(출처:Introduction to Applied Linear Algebra by Boyd and Vandenberghe, 2019)

$$\min_{\beta} \sum_{i=1}^{n} (y_i - \boldsymbol{x}_i^t \boldsymbol{\beta})^2 = \min_{\beta} (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^t (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) \tag{6}$$

위에서 $\boldsymbol{X}\boldsymbol{\beta}$는 행렬 $\boldsymbol{X}$의 열벡터로 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 벡터는 반응값 벡터 $\boldsymbol{y}$ 를 행렬 $\boldsymbol{X}$의 열벡터로 생성한 열공간에 사영(projection)하게 만드는 벡터이며

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

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

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

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