

4

직교성

4.1 벡터와 부분공간의 직교성

4.2 직선과 부분공간 위로의 사영

4.3 최소제곱근사

4.4 정규직교기저와 그람-슈미트 과정

4.5 행렬의 유사역행렬

4장에 대한 고찰 : 직교성

두 벡터가 직교한다는 것은 두 벡터의 내적이 0임을 의미한다. 즉 $v \cdot w = v^T w = 0$ 이다. 이 장에서는 지금까지 배운 내용을 **직교하는 부분공간**(orthogonal subspace), **직교기저**(orthogonal basis), **직교행렬**(orthogonal matrix)로 확장할 것이다. 두 부분공간의 벡터, 기저를 이루는 벡터, 그리고 행렬 Q 의 열벡터는 모두 서로 직교한다. 이 결과는 다음과 같이 직각삼각형의 변 v 와 w 에 대한 피타고라스 정리 $a^2 + b^2 = c^2$ 을 떠올리면 쉽게 이해할 수 있다.

직교벡터 $v^T w = 0$ 이고 $\|v\|^2 + \|w\|^2 = \|v + w\|^2$ 이다.

위 식의 우변을 전개하면 $(v + w)^T (v + w)$ 이다. $v^T w = w^T v = 0$ 이면 식 $(v + w)^T (v + w)$ 은 $v^T v + w^T w$ 가 된다.

부분공간 개념은 3장에서 $Ax = b$ 를 이해하기 위해 도입했으며, 이때 열공간과 영공간이 필요했다. 그다음 A^T 를 도입하면서 추가로 두 부분공간이 등장했다. 이 4가지 기본 부분공간은 행렬이 실제로 수행하는 작용을 명확히 보여준다.

행렬 A 와 벡터 x 의 곱 Ax 에 대해 생각해보자. 첫 번째 수준에서는 Ax 가 단순한 수치 연산에 해당한다. 두 번째 수준에서는 Ax 가 열들의 선형결합으로 표현된다. 세 번째 수준에서는 Ax 를 통해 부분공간의 구조가 드러난다. 그러나 뒤에서 나올 [그림 4.1]을 살펴보기 전까지는 전체적인 구조를 완전히 이해했다고 할 수 없다. 4가지 기본 부분공간은 서로 직교하는 관계이다.

영공간 $N(A)$ 는 행공간 $C(A^T)$ 와 직교하는 모든 벡터를 포함한다.

A^T 의 영공간 $N(A^T)$ 는 열공간 $C(A)$ 와 직교하는 모든 벡터를 포함한다.

식 $Ax = 0$ 은 x 가 행렬 A 의 각 행과 직교함을 의미하며, 식 $A^T y = 0$ 은 y 가 행렬 A 의 각 열과 직교함을 의미한다.

이 장에서 다룰 핵심 개념은 **사영**(projection)이다. 만약 벡터 b 가 행렬 A 의 열공간 바깥에 존재하면, 그 안에서 벡터 b 와 가장 가까운 점 p 를 찾아야 한다. 벡터 b 에서 p 까지의 직선은 오차벡터 e 를 나타내며, 오차벡터 e 는 행렬 A 의 열공간과 직교한다. **최소제곱 방정식**(least squares equation) $A^T A \hat{x} = A^T b$ 는 가장 가까운 점 $p = A \hat{x}$ 을 구하여 오차벡터 e 의 크기를 최소화한다. 방정식 $A^T A \hat{x} = A^T b$ 는 $Ax = b$ 의 해가 존재하지 않는 경우에, 최적의 $\hat{x}$ 을 제공한다. 최적의 $\hat{x}$ 에 대해 $\|A \hat{x} - b\|^2$ 이 최소가 된다. 이를 **최소제곱법**(least squares)이라 한다.

식 $A^T A \hat{x} = A^T b$ 는 $A^T A = I$ 일 때 간단히 풀린다. 이 경우 행렬 A 는 **정규직교**(orthonormal) 열벡터를 갖는다. 여기서 정규직교벡터란 서로 직교하며 크기가 1인 벡터를 말한다. 이러한 벡터의 조건은 자연적으로 잘 발생하지는 않지만 의도적으로 만들어낼 수 있다. 4.4절에서는 열벡터 $a_1, \dots, a_n$ 을 직교화하여 $q_i^T q_j = 0$ 을 만족하는 벡터 $q_1, \dots, q_n$ 을 구성하는 방법에 대해 알아볼 것이다. $q_1, \dots, q_n$ 을 열로 갖는 행렬 Q 는 $Q^T Q = I$ 와 $A = QR$ 을 만족한다. 여기서 R 은 상삼각행렬이다.

직교행렬은 행렬 계산에서 매우 유리하다. 두 직교행렬 Q_1 과 Q_2 를 곱하면 그 결과도 직교행렬이 된다. 여러 측면에서 $A = QR$ 분해는 $A = LU$ 분해보다 더 우수하며, 이 장에서 그 이유를 자세히 살펴볼 것이다.

4.1 벡터와 부분공간의 직교성

1. 두 벡터 v, w 가 $v^T w = 0$ 을 만족하면 서로 직교한다고 한다. 이때 $\|v\|^2 + \|w\|^2 = \|v + w\|^2$ 이 성립하며, 이는 $a^2 + b^2 = c^2$ 과 같은 피타고라스 정리에 해당한다.
2. 모든 $v \in V$ 와 $w \in W$ 에 대해 $v^T w = 0$ 이면 부분공간 V 와 W 가 서로 직교한다고 한다.
3. 행렬 A 의 행공간은 영공간 $N(A)$ 와 직교하며, 열공간은 A^T 의 영공간 $N(A^T)$ 와 직교한다.
4. 4가지 기본 부분공간의 차원에 대해 $r + (n - r) = n$ 과 $r + (m - r) = m$ 이 성립한다. 이런 관계를 만족하는 공간을 **직교여공간**(orthogonal complements)이라 한다.
5. $\mathbb{R}^n$ 에서 n 개의 벡터가 선형독립이면, 이 벡터들은 $\mathbb{R}^n$ 을 생성한다. 반대로 n 개의 벡터가 $\mathbb{R}^n$ 을 생성하면, 이 벡터들은 선형독립이다.

1장에서는 내적 $v^T w$ 를 두 벡터 v, w 의 사잇각과 연결했다. 사잇각이 90° 이면 $v^T w = 0$ 이 성립한다. 벡터 v, w 와 $v + w$ 는 직각삼각형을 형성한다. 직각삼각형의 변의 길이 a, b, c 에 대해 피타고라스 정리 $a^2 + b^2 = c^2$ 이 성립하는 것처럼, 벡터의 크기 $\|v\|, \|w\|, \|v + w\|$ 에 대해 다음과 같이 피타고라스 정리가 성립한다.

직교벡터 $v^T w = 0$ 이고 $\|v\|^2 + \|w\|^2 = \|v + w\|^2$ 이다.

앞 식의 우변 $\|v+w\|^2$ 은 $(v+w)^T(v+w) = v^T v + w^T w + v^T w + w^T v = v^T v + w^T w + 0 + 0$ 이 된다. 이러한 관계는 평면상에서도, n 차원 공간 $\mathbb{R}^n$ 에서도 성립한다.

직선 위의 모든 벡터와 평면 위에 모든 벡터가 서로 직교하면, 그 직선(1차원 부분공간)과 평면(2차원 부분공간)은 서로 직교한다. 이 결과는 3차원 이상의 공간에서도 마찬가지로 성립한다. 부분공간의 직교성은 임의의 행렬 A 의 행공간과 영공간을 통해 확인할 수 있다.

행렬 A 의 영공간은 A 의 행공간과 직교한다. $Ax = 0$ 을 살펴보자.

$$Ax = \begin{bmatrix} A \text{의 (1행)} \\ \vdots \\ A \text{의 (} m \text{행)} \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad \begin{array}{l} \leftarrow (1\text{행}) \cdot x = 0 \\ \leftarrow (m\text{행}) \cdot x = 0 \end{array} \quad (1)$$

행렬 A 의 각 행과 x 의 내적을 계산한 결과는 0이다. 따라서 행렬 A 의 행들의 모든 선형결합도 x 와 직교한다. 결국 전체 행공간 $C(A^T)$ 는 전체 영공간과 직교한다.

표기할 때는 열벡터를 기준으로 한다. 행렬 A 의 행벡터는 전치행렬 A^T 의 열벡터가 된다. 행렬 A 의 행공간의 벡터들은 이 열벡터들의 선형결합 $A^T y$ 로 표현된다. 부분공간이 서로 직교함을 보이는 간단한 증명은 다음과 같다.

$$\text{영공간의 벡터 } x \text{와 행공간의 벡터 } A^T y \text{가 직교} \quad x^T(A^T y) = (Ax)^T y = 0^T y = 0 \quad (2)$$

식 (1)과 같은 첫 번째 증명 방식은 직관적이다. 각 행벡터와 x 를 곱하면 0이 되고, 행의 모든 선형결합도 x 와 곱하면 0이 됨을 직접 확인할 수 있기 때문이다.

중요한 사실은 또 다른 한 쌍의 기본 부분공간이 존재한다는 것이다. 열공간 $C(A)$ 와 A^T 의 영공간 $N(A^T)$ 는 m 차원 공간 $\mathbb{R}^m$ 에서 서로 직교한다.

앞서 살펴본 원래 증명에 A 대신 A^T 를 대입해 적용할 수도 있다. A^T 의 행공간은 A 의 열공간과 같다. 따라서 (A^T 에 대해) ‘행공간과 영공간이 직교한다’는 성질은 ‘ A 의 열공간 $C(A)$ 와 A^T 의 영공간 $N(A^T)$ 가 직교한다’는 성질로 바뀐다.

EX 1 행렬 A 의 두 행은 A 의 영공간 안에 있는 벡터 x 와 직교한다.

$$\begin{array}{ll} \text{행공간} & Ax = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{array}{l} \text{내적 } 1-2+1=0 \\ \text{내적 } 1+0-1=0 \end{array} \\ \text{영공간} & \end{array}$$

또한 행렬 A 의 세 열벡터(또는 A^T 의 행벡터)는 $N(A^T)$ 안의 벡터 $y=0$ 과 직교한다.

$$\begin{array}{l} \text{열공간} \\ \text{좌영공간} \end{array} \quad A^T y = \begin{bmatrix} 1 & 1 \\ -2 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

이 예는 극단적인 경우로, 열공간 $C(A)$ 가 $\mathbb{R}^2$ 전체이고 A^T 의 영공간이 영벡터만 포함한다. 행공간과 영공간의 차원의 합은 $2+1=3$ 이며, 이 두 부분공간은 $\mathbb{R}^3 = \mathbb{R}^n$ 의 모든 벡터를 포함한다. 한편 열공간과 좌영공간(A^T 의 영공간)의 차원의 합은 $2+0=2$ 이며, 이 두 부분공간은 $\mathbb{R}^2 = \mathbb{R}^m$ 의 모든 벡터를 포함한다.

직교하는 임의의 두 부분공간 쌍에 대한 중요한 차원 조건이 다음과 같이 존재한다.

V 와 W 가 $\mathbb{R}^n$ 에서 직교하는 부분공간이면, $\dim V + \dim W \leq n$ 이 성립한다.

이 결과는 매우 명확하고 자연스러워 보인다. 예를 들어 방의 벽과 바닥을 생각해보자. 벽과 바닥은 $\mathbb{R}^3$ 에서 각각 2차원 부분공간이다. 두 부분공간의 차원의 합 $2+2=4$ 는 3보다 크므로, 두 부분공간은 직교하지 않는다. 일부 벡터는 두 공간에 동시에 포함되는데, 이는 벽과 바닥이 만나는 직선에 해당한다.

전체 공간을 구성하는 직교하는 두 부분공간의 쌍을 특별히 **직교여공간**이라 한다. V 의 직교여공간 $V^\perp$ 은 V 와 직교하는 모든 벡터를 포함한다. 선형대수학의 큰 구조에서 두 쌍의 기본 부분공간은 이러한 직교여공간 관계를 만족한다.

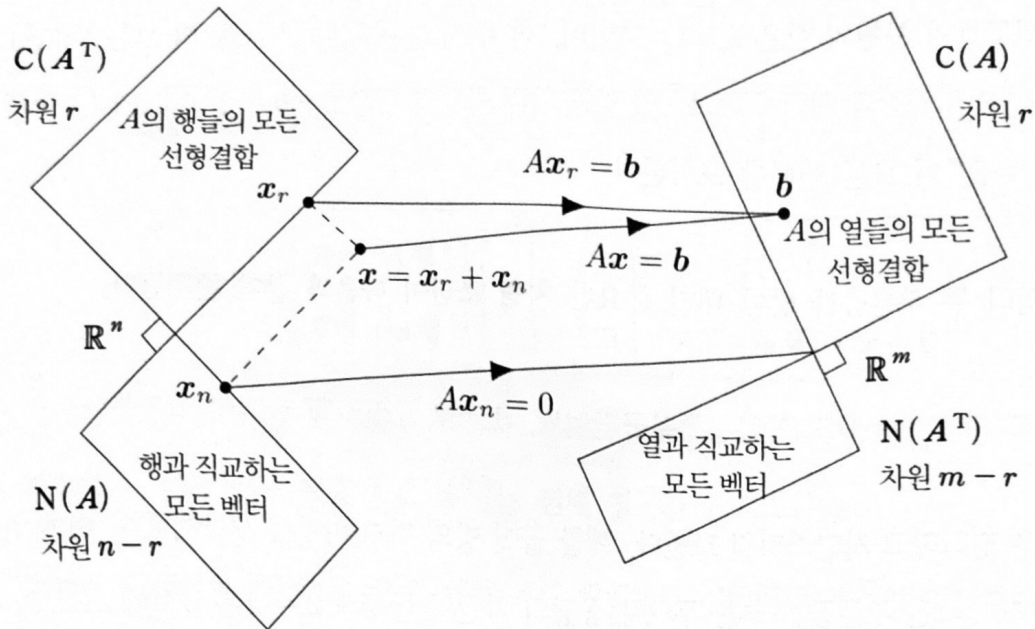
$$\begin{array}{l} \text{직교여공간} \\ \text{행공간과 영공간 : } r + (n - r) = n \\ \text{열공간과 좌영공간 : } r + (m - r) = m \end{array}$$

$\mathbb{R}^n$ 의 임의의 벡터 x 는 행공간의 벡터 x_{row} 와 영공간의 벡터 x_{null} 의 합 $x = x_{\text{row}} + x_{\text{null}}$ 로 나타낼 수 있다. 또한 $\mathbb{R}^m$ 의 임의의 벡터 y 는 열공간의 벡터 y_{col} 과 좌영공간(A^T 의 영공간)의 벡터 y_{null} 의 합 $y = y_{\text{col}} + y_{\text{null}}$ 로 나타낼 수 있다.

[정리] 선형대수학의 기본정리 2

- $\mathbb{R}^n$ 에서 $N(A)$ 는 행공간 $C(A^T)$ 의 직교여공간이다.
- $\mathbb{R}^m$ 에서 $N(A^T)$ 는 열공간 $C(A)$ 의 직교여공간이다.

3장의 [선형대수학의 기본정리 1]에서는 부분공간의 차원에 대해 다루었고, 이번 [선형대수학의 기본정리 2]에서는 그 부분공간들의 사잇각이 90° 라는 정보를 알려준다. $\mathbb{R}^n$ 안의 벡터 x 는 모두 행공간의 벡터 x_r 과 영공간의 벡터 x_n 으로 나눌 수 있다. $x = x_r + x_n$ 의 왼쪽에 행렬 A 를 곱할 때, [그림 4.1]은 $Ax_n = 0$, $Ax_r = Ax$ 임을 보여준다.



[그림 4.1] $Ax_r = b$ 는 열공간 안에 있고 $Ax_n = 0$ 이다. $Ax = b$ 의 완전해는 $x = (\text{하나의 } x_r) + (\text{임의의 } x_n)$ 이다.

그러면 $\|x\|^2 = \|x_r\|^2 + \|x_n\|^2$ 이고, $Ax = b$ 에 대한 **최소노름해**(minimum norm solution)는 (행공간의 $x = x_r$) + ($N(A)$ 의 $x_n = 0$)이다.

모든 벡터 Ax 는 열공간 안에 있다. 행렬 A 를 곱한다고 해서 할 수 있는 일은 없다. 열공간의 모든 벡터 b 는 행공간의 정확히 '하나의' 벡터 x_r 로부터 나온다. 이를 증명해보자. 만약 $Ax_r = Ax_r'$ 이면 차이 $x_r - x_r'$ 은 영공간 안에 있다. 또한 $x_r - x_r'$ 은 x_r 과 x_r' 을 포함한 행공간 안에 있다. 차이 $x_r - x_r'$ 은 영벡터여야 한다. 왜냐하면 영공간과 행공간이 직교하기 때문이다. 따라서 $x_r = x_r'$ 이다.

두 영공간을 제거하면 행렬 A 의 내부에 $r \times r$ 가역행렬이 숨어 있음을 알 수 있다. 행렬 A 는 행공간에서 열공간으로의 가역변환이다. 4.5절에서 살펴볼 유사역행렬 A^+ 는 열공간을 행공간으로 다시 변환한다.

EX 2 모든 랭크 r 행렬은 가역인 $r \times r$ 부분행렬을 포함한다. 다음 행렬 A 의 랭크는 2이다.

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 5 & 6 \\ 1 & 2 & 4 & 5 & 6 \end{bmatrix} \text{는 피벗 행과 피벗 열에서 } 2 \times 2 \text{ 부분행렬 } \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \text{를 포함한다.}$$

이제 임의의 행렬 A 에 대해 증명해보자. 부분행렬 C 는 r 개의 선형독립인 열을 갖는다(1장 참고). 그러면 C 와 C^T 의 랭크는 r 이다. 따라서 C^T 는 r 개의 선형독립인 열을 갖는다. 다시 말해 C 는 r 개의 선형독립인 행을 갖는다. 이를 통해 행렬 A 의 가역인 부분행렬을 찾는다.

다시 한 번 강조하자면, 직교하는 두 부분공간에 속하는 유일한 벡터는 영벡터이다.

부분공간으로부터의 기저 결합

기저는 벡터공간을 생성하며 선형독립인 벡터들의 집합이다. 일반적으로는 이와 같은 기저의 두 성질을 모두 확인해야 한다. 다만, 공간의 차원을 알 경우에는 다음과 같이 한 성질이 다른 성질을 암시한다. 모든 벡터는 정확히 1가지 방식으로 표현되는 기저벡터들의 선형결합이다.

- $\mathbb{R}^n$ 의 선형독립인 n 개의 벡터는 반드시 $\mathbb{R}^n$ 을 생성한다. 따라서 이 벡터들은 기저를 이룬다.
- $\mathbb{R}^n$ 을 생성하는 n 개의 벡터는 선형독립이어야 한다. 따라서 이 벡터들은 기저를 이룬다.

이 결과는 임의의 벡터공간에 대해 참이지만, 여기서는 $\mathbb{R}^n$ 위주로 생각한다. 어떤 벡터가 $n \times n$ 정사각행렬 A 의 열에 들어갈 때, 다음이 성립한다.

- 행렬 A 의 n 개의 열벡터가 선형독립이면 이 n 개의 열벡터들은 $\mathbb{R}^n$ 을 생성한다. 따라서 $Ax = b$ 의 해가 존재한다.
- n 개의 열벡터가 $\mathbb{R}^n$ 을 생성하면 이 n 개의 열벡터들은 선형독립이다. 따라서 $Ax = b$ 의 해는 유일하게 존재한다.
- 정사각행렬 A, B 에 대해 $AB = I$ 일 때 $BA = I$ 이다.

$Ax = b$ 에서 해의 유일성은 해가 존재함을 의미하고, 해의 존재성은 해가 유일함을 의미한다. 그러면 A 는 가역행렬이다. 자유변수가 없는 경우 해 x 는 유일하다. 따라서 n 개의 피벗 열이 있어야 한다. 그러면 역대입법으로 $Ax = b$ 의 해를 구할 수 있다(즉 해가 존재한다).

이번에는 반대 방향으로 시작하여, $Ax = b$ 의 해가 모든 벡터 b 에 대해 존재한다고 가정해보자. 소거를 진행하면 어떤 행도 영행이 되지 않는다. n 개의 피벗변수만 있고 자유변수는 없다. 따라서 행렬 A 의 영공간은 $x = 0$ 만 포함한다(즉 해가 유일하다).

행공간의 기저벡터 r 개와 영공간의 기저벡터 $(n-r)$ 개를 더하면 $r + (n-r) = n$ 개의 벡터가 된다. 이 n 개의 벡터는 선형독립이다.¹ 따라서 $\mathbb{R}^n$ 을 생성한다.

$\mathbb{R}^n$ 의 벡터 x 는 행공간의 벡터 x_r 과 영공간의 벡터 x_n 의 합 $x_r + x_n$ 이다. [그림 4.1]에서 분리된 $x_r + x_n$ 은 직교여공간의 핵심을 보여준다. 행공간과 영공간의 차원을 합하면 n 이 되고, 이는 모든 벡터에 대해 적용되는 사실이다.

EX ③ 행렬 $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ 에 대해 $x = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ 을 $x_r + x_n = \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ 로 분리해보자. 벡터 $(2, 4)$ 는 행렬 A 의 행공간 안에 있다. 직교벡터 $(2, -1)$ 은 행렬 A 의 영공간 안에 있다. 다음 절에서는 사영행렬 P 를 이용한 분리에 대해 살펴볼 것이다.

EX ④ S 를 9차원 공간 $\mathbb{R}^9$ 의 6차원 부분공간이라고 가정하면 다음이 성립한다.

- (a) S 와 직교하는 부분공간의 가능한 차원은 0, 1, 2, 3이다.
 - (b) S 의 직교여공간 $S^\perp$ 의 가능한 차원은 3이다.
 - (c) 행공간이 S 인 행렬 A 의 가능한 한 가장 작은 크기는 6×9 이다.
 - (d) 영공간이 $S^\perp$ 인 행렬 B 의 가능한 한 가장 작은 크기는 6×9 이다.
-

4.1 연습문제

1. 랭크가 1인 2×3 행렬을 구성하라. 그리고 [그림 4.1]의 각 부분공간에 벡터를 하나씩 작성하라(단, 영공간에는 2개를 작성하라). 어떤 벡터가 서로 직교하는가?

¹ 이를 증명해보자. n 개의 모든 벡터의 선형결합이 $x_r + x_n = 0$ 이면, 두 부분공간에서 $x_r = -x_n$ 이 된다. 따라서 $x_r = x_n = 0$ 이고, 행공간의 기저와 영공간의 기저에 대한 모든 계수는 0이어야 한다. 이 결과는 n 개의 모든 벡터가 선형독립임을 증명한다.

2. 랭크가 $r=2$ 인 3×2 행렬을 기준으로 [그림 4.1]을 다시 작성하라. 어떤 부분공간이 Z (영벡터만 포함)가 되는가? $\mathbb{R}^2$ 의 임의의 벡터 x 의 영공간 부분은 $x_n = \underline{\hspace{2cm}}$ 이다.
3. 다음 성질을 만족하는 행렬을 구성하라. 불가능한 경우에는 행렬을 구성할 수 없는 이유를 설명하라.
- (a) 열공간은 $\begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$ 를 포함하고, 영공간은 $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 을 포함한다.
- (b) 행공간은 $\begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$ 를 포함하고, 영공간은 $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 을 포함한다.
- (c) $Ax = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 의 해가 존재하고 $A^T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ 이다.
- (d) 행렬 A 의 모든 행은 모든 열과 직교한다(단, A 는 영행렬이 아니다).
- (e) 행렬의 모든 열을 더하면 성분이 모두 0인 열이 되고, 모든 행을 더하면 성분이 모두 1인 행이 된다.
4. $AB=0$ 이면 행렬 B 의 열은 행렬 A 의 $\underline{\hspace{2cm}}$ 안에 있다. 또 행렬 A 의 행은 행렬 B 의 $\underline{\hspace{2cm}}$ 안에 있다. $AB=0$ 인 경우, 왜 행렬 A, B 는 랭크가 2인 3×3 행렬이 될 수 없는가?
5. 다음 물음에 답하라.
- (a) $Ax=b$ 의 해가 존재하고 $A^T y=0$ 일 때, $(y^T x=0)$ 혹은 $(y^T b=0)$ 인가?
- (b) $A^T y=(1, 1, 1)$ 의 해가 존재하고 $Ax=0$ 이면 $(1, 1, 1)$ 은 행공간과 영공간에 대해 어떤 성질을 갖는가?
6. 해가 없는 임의의 연립방정식 $Ax=b$ 는 $0=1$ 형태가 될 수 있다. 다음 세 방정식에 y_1, y_2, y_3 를 각각 곱한 다음 모두 더했을 때 $0=1$ 형태가 되도록 하는 y_1, y_2, y_3 를 구하라. y_1, y_2, y_3 를 성분으로 갖는 벡터 y 는 어떤 부분공간 안에 있는가? 내적 $y^T b$ 가 1이므로 해 x 는 존재하지 않는다.

$$\begin{aligned} x+2y+2z &= 5 \\ 2x+2y+3z &= 5 \\ 3x+4y+5z &= 9 \end{aligned}$$

7. 해가 없는 임의의 연립방정식 $Ax = b$ 는 [문제 6]의 경우와 같다. m 개의 방정식에 $y_1, \dots, y_m$ 을 각각 곱한 다음 모두 더하면 $0 = 1$ 형태가 된다. 이 결과를 **프레드홀름 대안**(Fredholm's alternative)이라 한다(벡터 b 가 $C(A)$ 안에 없으면, b 의 일부는 $N(A^T)$ 안에 있다). 다음 둘 중 정확히 하나의 문제만 해를 갖는다.

$$Ax = b \quad \text{또는} \quad (y^T b = 1 \text{ 일 때, } A^T y = 0)$$

세 방정식 $x_1 - x_2 = 1, x_2 - x_3 = 1, x_1 - x_3 = 1$ 에 y_1, y_2, y_3 를 각각 곱한 다음 모두 더했을 때 $0 = 10$ 이 되도록 하는 y_1, y_2, y_3 를 구하라. 그러면 이 연립방정식은 해를 갖지 않는다.

8. [그림 4.1]에서 Ax , Ax 와 같음을 어떻게 알 수 있는가? 또 벡터 Ax 가 열공간 안에 있음을 어떻게 알 수 있는가? $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ 이고 $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 이면 x 은 무엇인가?
9. $A^T Ax = 0$ 이면 $Ax = 0$ 이다. 왜냐하면 Ax 는 A^T 의 영공간 안에 있으며, A 의 _____ 안에 있기 때문이다. 이 두 부분공간 사이에 어떤 관계가 성립하는가? 결론적으로 $Ax = 0$ 이고, 따라서 $A^T A$ 와 A 는 동일한 영공간을 갖는다.
10. A 가 대칭행렬일 때($A^T = A$), 다음 물음에 답하라.
- (a) 왜 열공간이 영공간과 직교하는가?
 - (b) $Ax = 0, Az = 5z$ 일 때, 어떤 부분공간이 '고유벡터(eigenvector)'² x, z 를 갖는가? 대칭행렬은 $x^T z = 0$ 과 같이 직교하는 고유벡터 x, z 를 갖는다.

11. [그림 4.1]을 그려서 행렬 $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix}$ 에 대한 각 부분공간을 나타내라.

12. $A = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, x = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ 에 대해 적절한 x', x'' 을 찾아서 [그림 4.1]을 그려라.

※ [문제 13~23]은 직교하는 부분공간에 관한 문제이다.

13. 두 부분공간 V, W 의 기저벡터를 각각 행렬 V 와 W 의 열에 넣는다. 두 부분공간 V, W 가 직교하는 부분공간인지를 확인하는 판정법이 ' $V^T W = 0$ (영행렬)'인 이유를 설명하라. 이는 v, w 가 직교벡터인 경우 $v^T w = 0$ 을 만족한다는 사실과 유사하다.

2 (편집자) 고유벡터에 대해서는 6장에서 자세히 설명한다.

14. 바닥 V 와 벽 W 는 직교하는 부분공간이 아니다. 왜냐하면 두 부분공간이 서로 만나는 선을 따라 영벡터가 아닌 벡터를 공유하기 때문이다. $\mathbb{R}^3$ 의 어떤 두 평면 V, W 도 서로 직교할 수 없다. 다음 두 행렬의 열공간에 모두 속하는 벡터를 구하라. 이 벡터들은 벡터 Ax 와 $B\hat{x}$ 으로 표현될 것이다. 3×4 행렬 $[A \ B]$ 를 이용해 생각해보자.

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 4 \\ 6 & 3 \\ 5 & 1 \end{bmatrix}$$

15. [문제 14]를 p 차원 부분공간 V 와 q 차원 부분공간 W 로 확장하자. V 와 W 의 교집합의 원소가 영벡터가 아닌 벡터라는 사실을 보장하는 $(p+q)$ 에 대한 부등식을 구하라. 두 부분공간 V, W 는 서로 직교할 수 없다.
16. 식 (2)의 행렬 용어를 이용해 $N(A^T)$ 의 벡터 y 가 열공간의 모든 벡터 Ax 와 직교함을 증명하라. $A^T y = 0$ 에서 시작하라.
17. S 가 영벡터만 포함하는 $\mathbb{R}^3$ 의 부분공간일 때, $S^\perp$ 은 무엇인가? S 가 벡터 $(1, 1, 1)$ 로 생성되는 경우, $S^\perp$ 은 무엇인가? S 가 두 벡터 $(1, 1, 1), (1, 1, -1)$ 로 생성되는 경우, $S^\perp$ 의 기저는 무엇인가?
18. S 는 두 벡터 $(1, 5, 1), (2, 2, 2)$ 만 포함한다(부분공간이 아님). 그러면 $S^\perp$ 은 어떤 행렬 A 의 영공간인가? S 는 부분공간이 아니더라도 $S^\perp$ 은 부분공간이다.
19. L 은 $\mathbb{R}^3$ 안의 1차원 부분공간(직선)이다. L 의 직교여공간 $L^\perp$ 은 L 과 직교하는 __차원 부분공간이다. 그러면 $(L^\perp)^\perp$ 은 $L^\perp$ 과 직교하는 __차원 부분공간이다. 사실 $(L^\perp)^\perp$ 은 __와(과) 동일하다.
20. V 를 $\mathbb{R}^4$ 전체 공간이라 가정하자. 그러면 $V^\perp$ 은 단지 _____만 포함한다. 따라서 $(V^\perp)^\perp$ 은 _____이다. 따라서 $(V^\perp)^\perp$ 은 __와(과) 동일하다.
21. S 는 두 벡터 $(1, 2, 2, 3), (1, 3, 3, 2)$ 로 생성된다. $S^\perp$ 을 생성하는 두 벡터를 구하라. 이 문제는 어떤 행렬 A 에 대해 $Ax = 0$ 을 푸는 것과 같은가?
22. P 가 $\mathbb{R}^4$ 안에서 $x_1 + x_2 + x_3 + x_4 = 0$ 을 만족하는 벡터들의 평면일 때, $P^\perp$ 의 기저를 구하라. 또한 P 를 영공간으로 갖는 행렬을 구하라.
23. 부분공간 S 가 부분공간 V 에 포함될 때, $S^\perp$ 이 $V^\perp$ 을 포함하는 이유를 설명하라.

※ [문제 24~28]은 직교하는 열과 행에 관한 문제이다.

24. $n \times n$ 행렬 A 가 가역행렬이라 가정하자($AA^{-1}=I$). 그러면 A^{-1} 의 첫 번째 열은 행렬 A 의 어떤 행이 생성하는 공간과 직교하는가?
25. 행렬 A 의 열들이 서로 직교하는 단위벡터일 때, $A^T A$ 를 구하라.
26. 모든 열이 서로 직교하고 어떤 성분도 0이 아닌 3×3 행렬 A 를 만들어라. 그리고 $A^T A$ 를 계산하라. $A^T A$ 는 왜 대각행렬이 되는가?
27. 두 직선 $3x + y = b_1$ 과 $6x + 2y = b_2$ 는 서로 어떤 관계인가? 만약 _____이면 두 직선은 서로 일치한다. 이 경우 벡터 (b_1, b_2) 는 어떤 벡터와 직교하는가? 두 직선의 방정식으로 만들어지는 행렬의 영공간은 직선 $3x + y = \underline{\hspace{1cm}}$ 이다. 이 영공간의 특정 벡터 하나는 _____이다.
28. 다음 명제가 거짓인 이유를 제시하라.
 - (a) $(1, 1, 1)$ 은 $(1, 1, -2)$ 와 직교하므로, 두 평면 $x + y + z = 0$ 과 $x + y - 2z = 0$ 은 직교하는 부분공간이다.
 - (b) $(1, 1, 0, 0, 0)$ 과 $(0, 0, 0, 1, 1)$ 에 의해 생성되는 부분공간은 $(1, -1, 0, 0, 0)$ 과 $(2, -2, 3, 4, -4)$ 에 의해 생성되는 부분공간의 직교여공간이다.
 - (c) 영벡터에서만 만나는 두 부분공간은 서로 직교한다.
29. 벡터 $v = (1, 2, 3)$ 이 행공간과 열공간 안에 있는 행렬 A 를 구하라. 또 벡터 v 가 영공간과 열공간 안에 있는 행렬 B 를 구하라. 벡터 v 를 공유할 수 없는 부분공간 쌍은 무엇인가?
30. A 는 3×4 행렬, B 는 4×5 행렬이고 $AB = 0$ 이라 하자. 그러면 $N(A)$ 는 $C(B)$ 를 포함한다. $N(A)$ 와 $C(B)$ 의 차원을 이용해 $\text{rank}(A) + \text{rank}(B) \leq 4$ 임을 증명하라.
31. 명령어 $N = \text{null}(A)$ 가 행렬 A 의 영공간의 기저를 만든다고 가정하자. 그러면 명령어 $B = \text{null}(N^T)$ 는 행렬 A 의 _____의 기저를 생성한다.
32. $\mathbb{R}^2$ 의 영벡터가 아닌 네 벡터 r, n, c, l 이 2×2 행렬 A 의 4가지 기본 부분공간 $C(A^T)$, $N(A)$, $C(A)$, $N(A^T)$ 의 기저벡터가 되기 위한 조건은 무엇인가?
33. $\mathbb{R}^4$ 의 영벡터가 아닌 8개의 벡터 $r_1, r_2, n_1, n_2, c_1, c_2, l_1, l_2$ 가 4×4 행렬의 4가지 기본 부분공간의 기저벡터가 되기 위한 조건은 무엇인가?

4.2 직선과 부분공간 위로의 사영

1. 벡터 b 의 벡터 a 방향 직선 위로의 사영은 직선에서 벡터 b 와 가장 가까운 점 $p = a \frac{a^T b}{a^T a}$ 이다.
2. 오차벡터 $e = b - p$ 는 벡터 a 와 직교한다. 직각삼각형 bpe 는 $\|p\|^2 + \|e\|^2 = \|b\|^2$ 을 만족한다.
3. 벡터 b 의 부분공간 S 위로의 사영은 부분공간 S 안에서 벡터 b 와 가장 가까운 벡터 p 이다. $b - p$ 는 S 와 직교한다.
4. 행렬 A 가 선형독립인 열을 가질 때, $A^T A$ 는 가역인 대칭행렬이다($N(A^T A) = N(A)$).
5. 벡터 b 의 열공간 $C(A)$ 위로의 사영은 벡터 $p = A(A^T A)^{-1} A^T b$ 이다.
6. $C(A)$ 위로의 사영행렬(projection matrix)은 $P = A(A^T A)^{-1} A^T$ 이다. 그리고 $p = Pb$, $P^2 = P = P^T$ 이다.

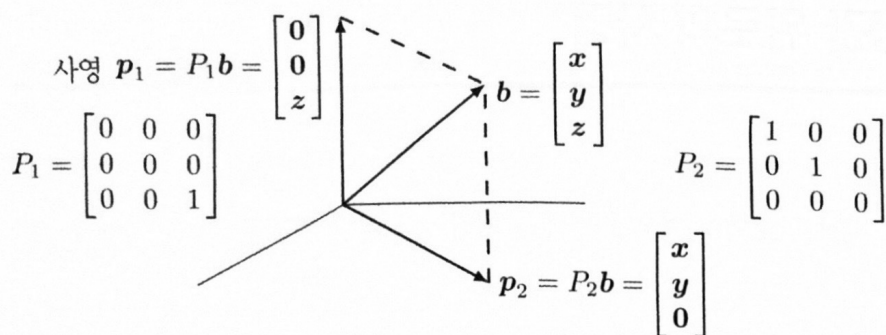
부분공간 S 위로의 사영은 쉽게 시각화할 수 있다. 벡터 b 는 S 안에서 가장 가까운 점 p 로 사영된다. 그러면 오차벡터 $e = b - p$ 는 부분공간 S 와 직교한다. 만약 행렬 A 가 선형독립인 열을 갖는다면, 벡터 b 의 $C(A)$ 위로의 사영은 $p = A(A^T A)^{-1} A^T b$ 이다. 여기서 사영행렬 $P = A(A^T A)^{-1} A^T$ 는 대칭행렬이다. 사영행렬의 특별한 성질은 $P^2 = P$ 이다. 즉 두 번째 사영은 p 를 p 로 사영하므로 아무것도 변하지 않는다.

특별히 좋은 성질을 가진 부분공간의 경우, 사영행렬 P 를 확인할 수 있다. 다음 질문을 통해 확인해보자.

- ① 벡터 $b = (2, 3, 4)$ 의 z -축과 xy -평면 위로의 사영은 무엇인가?
- ② 어떤 행렬 P_1, P_2 가 직선과 평면 위로의 사영을 만드는가?

벡터 b 가 직선 위로 사영되면, 사영 p 는 벡터 b 의 해당 직선 방향 부분이 된다. 벡터 b 가 평면 위로 사영되면, 사영 p 는 벡터 b 의 해당 평면 방향 부분이 된다. 사영 p 는 Pb 이다.

z -축 위로의 사영을 p_1 이라 하고, xy -평면 위로의 사영을 p_2 라 하자. 벡터 $b = (2, 3, 4)$ 를 z -축 위로 사영하면 $p_1 = (0, 0, 4)$ 이다. 또한 벡터 $b = (2, 3, 4)$ 를 xy -평면 위로 사영하면 $p_2 = (2, 3, 0)$ 이다. p_1 이 벡터 b 의 z -축 방향 부분이고, p_2 가 벡터 b 의 xy -평면 방향 부분이다 ([그림 4.2] 참고).



[그림 4.2] $P_1 + P_2 = I$, $P_1 P_2 = 0$ 을 만족하는 사영 $p_1 = P_1 b$ 와 $p_2 = P_2 b$

사영행렬 P_1 , P_2 는 3×3 행렬이다. 사영행렬과 3개의 성분을 가진 벡터 b 를 곱하면 3개의 성분을 가진 사영 p 가 만들어진다. 직선 위로의 사영행렬은 랭크가 1이고, 평면 위로의 사영행렬은 랭크가 2이다.

$$z\text{-축 위로의 사영행렬} : P_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad xy\text{-평면 위로의 사영행렬} : P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

사영행렬 P_1 은 모든 벡터의 z 축 성분을 뽑아낸다. 사영행렬 P_2 는 모든 벡터의 x 축 성분과 y 축 성분을 뽑아낸다. 벡터 b 의 사영 p_1 과 p_2 를 구하기 위해서는, 다음과 같이 b 에 각각 사영행렬 P_1 과 P_2 를 곱하면 된다(참고로 벡터는 소문자 p 로, 행렬은 대문자 P 로 표기한다).

$$p_1 = P_1 b = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}, \quad p_2 = P_2 b = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

이 경우 사영 p_1 과 p_2 는 서로 직교한다. xy -평면과 z 축은 방의 바닥면과 위쪽으로 향하는 벽의 직선과 같이 서로 직교하는 부분공간이다.

직선과 평면은 단순히 직교하는 것을 넘어, 서로 직교여공간이다. 두 공간(직선, 평면)의 차원을 더하면 $1 + 2 = 3$ 이다. 전체 공간 $\mathbb{R}^3$ 의 모든 벡터 b 는 두 부분공간에서의 벡터 b 의 두 부분의 합이다. 사영 $(0, 0, z)$, $(x, y, 0)$ 은 정확히 벡터 b 의 두 부분이다.

벡터들은 $p_1 + p_2 = b$ 를 만족하고, 행렬들은 $P_1 + P_2 = I$, $P_1 P_2 = 0$ 을 만족한다. (1)

이 예를 통해 우리의 목표를 달성했다. 어떤 직선이든, 어떤 평면이든, $\mathbb{R}^m$ 안의 어떤 n 차원 부분 공간이든 동일한 목표를 갖는다. 여기서 목표는 각 부분공간에서 부분 p 를 찾고, 부분 $p = Pb$ 를 생성하는 사영행렬 P 를 찾는 것이다. $\mathbb{R}^m$ 의 모든 부분공간은 해당하는 $m \times m$ 사영행렬 P 를 포함한다.

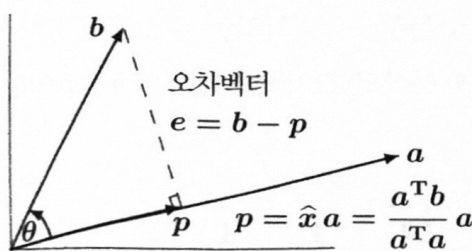
부분공간을 설명하는 가장 좋은 방법은 기저를 제시하는 것이다. 기저벡터들은 행렬 A 의 열에 넣는다. 이제 행렬 A 의 열공간 위로 사영한다. z 축은 3×1 행렬 A_1 의 열공간이다. xy -평면은 행렬 A_2 의 열공간이다. 이 평면은 행렬 A_3 의 열공간이기도 하다(부분공간의 기저는 많다). 따라서 $p_2 = p_3$, $P_2 = P_3$ 이다.

$$A_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 0 & 0 \end{bmatrix}$$

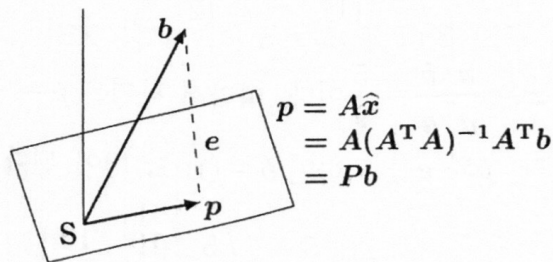
이제 생각해볼 문제는 $m \times n$ 행렬의 열공간 위로 임의의 벡터 b 를 사영하는 것이다. 우선 직선 (1차원)으로 시작하자. 이 경우 행렬 A 는 오직 하나의 열만 갖는다. 이 열을 a 라 하자.

직선 위로의 사영

$a = (a_1, \dots, a_m)$ 방향의 직선은 원점을 지난다. 이 직선 위에서 벡터 $b = (b_1, \dots, b_m)$ 과 가장 가까운 점 p 를 찾고자 한다. 사영의 핵심은 직교성이다. b 에서 p 까지의 직선은 벡터 a 와 직교한다. 이 직선은 [그림 4.3(a)]에서 점선으로 나타낸 오차벡터 $e = b - p$ 이다. 이제 p 를 유도해보자.



(a) 직선 위로의 사영



(b) S 위로의 사영

[그림 4.3] 직선과 $S(A$ 의 열공간) 위로의 벡터 b 의 사영 p

사영 p 는 a 의 상수배가 될 것이다. $p = \hat{x}a$ 는 'x 헷(hat)과 a 의 곱'으로 읽는다. $\hat{x}$ 을 계산하면 p 를 알 수 있다. 그런 다음 p 를 구하는 공식에서 사영행렬 P 를 구한다. 이 3단계로부터 모든 사영행렬을 찾는다($\hat{x}$ 을 구하고, 벡터 p 를 구한 다음, 행렬 P 를 구한다).

점선 $b - p$ 는 '오차벡터' $e = b - \hat{x}a$ 이다. $e = b - \hat{x}a$ 는 a 와 직교한다. 이 사실을 이용해 $\hat{x}$ 을 구해보자. $b - \hat{x}a$ 와 a 가 직교한다는 사실을 이용한다(즉 두 벡터의 내적이 0이다).

b 를 a 위로 사영

이때 오차벡터는 $e = b - \hat{x}a$

$$a \cdot (b - \hat{x}a) = 0 \text{ 또는 } a \cdot b - \hat{x}a \cdot a = 0$$

$$\hat{x} = \frac{a \cdot b}{a \cdot a} = \frac{a^T b}{a^T a} \quad (2)$$

$a^T b$ 는 $a \cdot b$ 와 같다. 다만 $a^T b$ 형태는 행렬에도 적용할 수 있으므로 더 유용하다. 공식 $\hat{x} = \frac{a^T b}{a^T a}$ 를 통해 사영 $p = \hat{x}a$ 를 알 수 있다.

벡터 b 의 벡터 a 방향 직선 위로의 사영은 $p = \hat{x}a = \frac{a^T b}{a^T a}a$ 이다.

• 특별한 경우 ① : $b = a$ 이면 $\hat{x} = 1$ 이다. 이때 벡터 a 의 a 위로의 사영은 자기 자신이다 ($Pa = a$).

• 특별한 경우 ② : b 와 a 가 직교하면 $a^T b = 0$ 이다. 이때 사영은 $p = 0$ 이다.

EX ① $b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 을 $a = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ 위로 사영하여, [그림 4.3]에서 $p = \hat{x}a$ 를 찾아보자.

$\hat{x}$ 은 $\frac{a^T b}{a^T a} = \frac{5}{9}$ 이다. 따라서 사영은 $p = \frac{5}{9}a$ 이다. b 와 p 사이의 오차벡터는 $e = b - p$ 이다. p 와 e 를 더하면 $b = (1, 1, 1)$ 이 된다.

$$p = \frac{5}{9}a = \left(\frac{5}{9}, \frac{10}{9}, \frac{10}{9}\right), \quad e = b - p = \left(\frac{4}{9}, -\frac{1}{9}, -\frac{1}{9}\right)$$

오차벡터 e 는 $a = (1, 2, 2)$ 와 직교한다. 즉 $e^T a = \frac{4}{9} - \frac{2}{9} - \frac{2}{9} = 0$ 이다.

b, p, e 를 세 변으로 갖는 직각삼각형을 생각해보자. 벡터 b 는 직선 방향 부분 p 와 이 직선과 직교하는 부분 e 로 분할된다. 두 부분 p 와 e 의 크기는 각각 $\|p\| = \|b\|\cos\theta$ 와 $\|e\| = \|b\|\sin\theta$ 이다. 삼각법에 따라 내적에 대해 다음이 성립한다.

$$p = \frac{a^T b}{a^T a}a \text{의 길이는 } \|p\| = \frac{\|a\|\|b\|\cos\theta}{\|a\|^2}\|a\| = \|b\|\cos\theta \text{이다.} \quad (3)$$

내적을 이용한 계산은 $\cos\theta$ 와 b 의 길이를 이용해 $\|p\| = \|b\|\cos\theta$ 를 계산하는 것보다 훨씬 간단하다. 실제 수치를 보면 $\cos\theta = \frac{5}{3\sqrt{3}}$, $\|b\| = \sqrt{3}$ 에는 제곱근이 있지만, 사영 $p = \frac{5a}{9}$ 에는 제곱근이 없다. 따라서 $\frac{a^T b}{a^T a}$ 를 이용하면 $\frac{5}{9}$ 라는 결론을 쉽게 얻을 수 있다.

이제 사영행렬이 등장한다. p 를 구하는 공식에서 b 에 어떤 행렬을 곱해야 하는가? $\hat{x}$ 을 a 의 오른 쪽으로 하면 사영행렬을 더 정확히 확인할 수 있다.

$$\text{사영행렬 } P \quad \text{사영행렬이 } P = \frac{aa^T}{a^T a} \text{ 일 때, 사영은 } p = a\hat{x} = a \frac{a^T b}{a^T a} = Pb$$

사영행렬 P 는 열과 행의 곱이다. 여기서 열은 a 이고, 행은 a^T 이다. 이렇게 둘을 곱한 결과를 수 $a^T a$ 로 나눈다. 사영행렬의 크기는 $m \times m$ 이지만 랭크는 1이다. 즉 1차원 부분공간인 직선 위로 사영하는 것이다. 이 직선은 사영행렬 P 의 열공간이다.

EX 2 벡터 $a = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ 방향의 직선 위로 사영하는 사영행렬 $P = \frac{aa^T}{a^T a}$ 를 구해보자. 우선 열 a 와 행 a^T 를 곱한 결과를 $a^T a = 9$ 로 나눈다.

$$\text{사영행렬} \quad P = \frac{aa^T}{a^T a} = \frac{1}{9} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix}$$

사영행렬 P 는 임의의 벡터 b 를 a 위로 사영한다. [EX 1]의 벡터 $b = (1, 1, 1)$ 에 대해 $p = Pb$ 임을 확인하면 다음과 같다.

$$p = Pb = \frac{1}{9} \begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 5 \\ 10 \\ 10 \end{bmatrix}$$

이때 오차벡터는 $e = b - p = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{9} \begin{bmatrix} 5 \\ 10 \\ 10 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 4 \\ -1 \\ -1 \end{bmatrix}$ 이다. 벡터 a 가 2배가 되어도 사영행렬 P 는 그대로 유지된다. 즉 사영행렬 P 는 벡터를 여전히 동일한 직선 위로 사영한다. 사영행렬의 제곱 P^2 은 P 와 같다. 왜냐하면 두 번 사영해도 아무것도 변하지 않기 때문이다 ($P^2 = P$). 행렬 P 의 주대각성분의 합은 $\frac{1}{9}(1 + 4 + 4) = 1$ 이다. 그리고 $e^T p = 0$ 이다.

항등행렬 I 에 대해 행렬 $I - P$ 도 사영행렬이어야 한다. 이 행렬은 삼각형의 다른 변 e 를 생성한다. P 가 하나의 부분공간 위로 사영할 때, $I - P$ 는 그 부분공간과 직교하는 부분공간 위로 사영한다. 즉 $I - P$ 는 a 와 직교하는 평면(직교여공간) 위로 사영한다.

이제 직선 위로의 사영을 넘어, $\mathbb{R}^m$ 의 n 차원 부분공간 위로의 사영에 대해 알아보자. 중요한 공식은 식 (5), (6), (7)에 등장한다. 이 3가지 식은 기본적으로 꼭 기억해두어야 한다.

부분공간 위로의 사영

$\mathbb{R}^m$ 안의 n 개의 벡터 $a_1, \dots, a_n$ 으로 시작한다. $a_1, \dots, a_n$ 은 선형독립이라고 가정하자. 여기서 문제 하나를 제시한다. “주어진 벡터 b 와 가장 가까운 선형결합 $p = \hat{x}_1 a_1 + \dots + \hat{x}_n a_n$ 을 구하라.” 우선 $\mathbb{R}^m$ 안의 벡터 b 를 $a_1, \dots, a_n$ 으로 생성되는 n 차원 부분공간 위로 사영한다.

$n = 1$ (하나의 벡터 a_1)인 경우, 위와 같은 사영은 직선 위로의 사영이다. 그리고 이 직선은 단 하나의 열만 갖는 행렬 A 의 열공간이다. 일반적으로 행렬 A 는 n 개의 열 $a_1, \dots, a_n$ 을 갖는다.

$\mathbb{R}^m$ 안에서의 선형결합은 열공간의 벡터 Ax 이다. 우리는 벡터 b 와 가장 가까운 선형결합 $p = A\hat{x}$ (사영)을 찾고자 한다. $\hat{x}$ 위의 $\hat{}$ (hat) 표기는 열공간 안에서 가장 가까운 벡터를 제공하는 최적의 선택임을 의미한다. $n = 1$ 인 경우, 그 최적의 선택은 $\hat{x} = \frac{a^T b}{a^T a}$ 이다. 이제 $n > 1$ 인 경우에 대한 최적의 선택 $\hat{x} = (\hat{x}_1, \dots, \hat{x}_n)$ 을 찾아보자.

앞서 ‘직선 위로의 사영’에서 살펴본 3단계로 n 차원 부분공간 위로의 사영을 계산하고자 한다. 즉 (S) 안에서 벡터 $\hat{x}$ 을 구하고, (C) 안에서 사영 $p = A\hat{x}$ 을 구한 다음, (T) 안에서 사영행렬 P 를 구해보자.

핵심은 기하학에 있다. [그림 4.3]의 점선은 벡터 b 로부터 부분공간 안에서 벡터 b 와 가장 가까운 지점인 $p = A\hat{x}$ 을 지난다. 오차벡터 $b - A\hat{x}$ 은 부분공간과 직교한다. 따라서 오차벡터 $b - A\hat{x}$ 은 부분공간의 모든 벡터 $a_1, \dots, a_n$ 과 직교한다. 이 n 개의 벡터가 이루는 직각에 대한 식은 다음과 같이 $\hat{x}$ 에 대한 n 개의 방정식을 유도한다.

$$\begin{array}{l} a_1^T (b - A\hat{x}) = 0 \\ \vdots \\ a_n^T (b - A\hat{x}) = 0 \end{array} \quad \text{또는} \quad \begin{bmatrix} -a_1^T - \\ \vdots \\ -a_n^T - \end{bmatrix} \begin{bmatrix} b - A\hat{x} \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix} \quad \text{또는} \quad A^T A \hat{x} = A^T b \quad (4)$$

a_i^T 들을 행으로 갖는 행렬은 A^T 이다. n 개의 방정식은 정확히 $A^T(b - A\hat{x}) = 0$ 이다. $A^T(b - A\hat{x}) = 0$ 을 잘 알려진 형태로 다시 작성하면 $A^T A \hat{x} = A^T b$ 이다. 이는 $\hat{x}$ 에 대한 방정식이며 계수행렬이 $A^T A$ 이다. 이제 $\hat{x}$, p , P 를 순서대로 구할 수 있다.

다음 세 방정식 (5), (6), (7)은 사영의 핵심이다.

b 와 가장 가까운 선형결합 $p = \hat{x}_1 a_1 + \dots + \hat{x}_n a_n = A\hat{x}$ 은 $\hat{x}$ 을 구하면 알 수 있다.

$$(n \times 1 \text{ 크기의}) \hat{x} \text{ 구하기} \quad A^T(b - A\hat{x}) = 0 \quad \text{또는} \quad A^T A \hat{x} = A^T b \quad (5)$$

대칭행렬 $A^T A$ 의 크기는 $n \times n$ 이다. $a_1, \dots, a_n$ 이 선형독립이면 $A^T A$ 는 가역행렬이다. 따라서 $A^T A \hat{x} = A^T b$ 의 해는 $\hat{x} = (A^T A)^{-1} A^T b$ 이다. 벡터 b 의 부분공간 위로의 사영은 p 이다.

$$(m \times 1 \text{ 크기의}) p \text{ 구하기} \quad p = A\hat{x} = A(A^T A)^{-1} A^T b \quad (6)$$

사영행렬 P 는 식 (6)에서 벡터 b 에 곱해진 행렬이다.

$$(m \times m \text{ 크기의}) P \text{ 구하기} \quad P = A(A^T A)^{-1} A^T \quad (7)$$

행렬 A 의 열이 하나만 있는 경우, 직선 위로의 사영과 비교해보자($A^T A = a^T a$ 의 크기는 1×1 이다).

$$n = 1 \text{인 경우} \quad \hat{x} = \frac{a^T b}{a^T a}, \quad p = a \frac{a^T b}{a^T a}, \quad P = \frac{a a^T}{a^T a} \quad (8)$$

식 (8)은 식 (5), (6), (7)과 동일하다. 수 $a^T a$ 는 행렬 $A^T A$ 가 된다. 각 계산에서 수 $a^T a$ 인 경우에는 그 수로 나누고, 행렬 $A^T A$ 인 경우에는 역행렬을 구한다. 열 $a_1, \dots, a_n$ 가 선형독립이면 $A^T A$ 는 가역행렬이다.

핵심은 정규방정식(normal equation) $A^T(b - A\hat{x}) = 0$ 을 해결하는 단계이다. 여기서 우리는 기하적 사실(오차벡터 e 는 각 벡터 a 와 직교)을 이용했다. 선형대수학은 다음과 같이 ‘정규방정식’을 매우 빠르고 아름다운 방법으로 도출한다.

❶ 부분공간 $C(A)$ 는 행렬 A 의 열공간이다.

❷ 오차벡터 $e = b - A\hat{x}$ 은 열공간과 직교하는 부분공간 $N(A^T)$ 안에 있다.

위 내용은 $A^T(b - A\hat{x}) = 0$ 을 의미한다. 좌영공간 $N(A^T)$ 는 사영에서 아주 중요하다. 좌영공간 $N(A^T)$ 에는 오차벡터 $e = b - A\hat{x}$ 이 포함된다. 벡터 b 는 사영 p 와 오차벡터 $e = b - p$ 로 분해된다. 사영은 변 p, e, b 로 이루어진 직각삼각형을 생성한다.

EX 3 $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$ 이고 $b = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$ 일 때, $\hat{x}$, p , P 를 구해보자. 우선 제곱행렬 $A^T A$ 와 벡터 $A^T b$ 를 구하면 다음과 같다.

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}, \quad A^T b = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

이제 정규방정식 $A^T A \hat{x} = A^T b$ 를 풀어서 $\hat{x}$ 을 구하면 다음과 같다.

$$\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix} \Rightarrow \hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix} \quad (9)$$

선형결합 $p = A\hat{x}$ 은 벡터 b 를 행렬 A 의 열공간 위로 사영한 것이다.

$$p = 5 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 3 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} \Rightarrow \text{오차벡터 } e = b - p = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad (10)$$

위 계산을 통해 2가지 사실을 확인해보자. 첫째, 오차벡터 $e = (1, -2, 1)$ 은 두 열벡터 $(1, 1, 1)$, $(0, 1, 2)$ 와 직교한다. 둘째, 행렬 P 과 벡터 $b = (6, 0, 0)$ 의 곱은 $p = (5, 2, -1)$ 이다. 즉 P 를 구하면 특정 벡터 b 를 곱함으로써 곧바로 p 를 구할 수 있다.

사영행렬은 $P = A(A^T A)^{-1} A^T$ 이다. $A^T A$ 의 행렬식은 $15 - 9 = 6$ 이다. 그러면 2×2 행렬 $A^T A$ 의 역행렬은 쉽게 구할 수 있다. 따라서 $A(A^T A)^{-1} A^T$ 를 계산하면 다음과 같이 사영행렬 P 를 구할 수 있다.

$$(A^T A)^{-1} = \frac{1}{6} \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix}, \quad P = \frac{1}{6} \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix} \quad (11)$$

두 번 사영한다고 해서 첫 번째 사영의 결과가 바뀌지 않으므로 $P^2 = P$ 여야 한다.

단, 행렬 $P = A(A^T A)^{-1} A^T$ 를 계산할 때 주의할 점이 있다. $(A^T A)^{-1}$ 를 A^{-1} 와 $(A^T)^{-1}$ 의 곱으로 나누어 생각할 수 있다. 만약 이 실수를 하고 P 를 구하는 식에 대입한다면 $P = A A^{-1} (A^T)^{-1} A^T$ 가 된다. 그러면 모든 항이 상쇄되어 $P = I$ 가 된다. 이제 왜 이런 잘못된 결과가 나왔는지 알아보자.

A 는 직사각행렬이다. 즉 A 의 역행렬은 없다. $m > n$ 인 경우에는 A^{-1} 가 존재하지 않으므로 $(A^T A)^{-1}$ 를 A^{-1} 와 $(A^T)^{-1}$ 로 나눌 수 없다.

경험상 직사각행렬에 관한 문제는 항상 $A^T A$ 로 이어진다. $m \times n$ 행렬 A 의 열이 선형독립일 때, $A^T A$ 는 가역행렬이다. 이 결과는 매우 중요하므로 다음과 같이 증명하고자 한다.

$A^T A$ 가 가역행렬일 필요충분조건은 행렬 A 의 열이 선형독립인 것이다.

증명

$A^T A$ 는 $n \times n$ 정사각행렬이다. 행렬 A 의 열이 선형독립일 때, A 의 영공간은 오직 영벡터만 포함한다. 이때 $A^T A$ 가 행렬 A 와 동일한 영공간을 가지면 $A^T A$ 는 가역행렬이 된다. 이에 따라 모든 행렬 A 에 대해 $A^T A$ 가 행렬 A 와 동일한 영공간을 가짐을 보임으로써 주어진 명제를 증명할 것이다.

A 를 임의의 행렬이라 하자. x 가 행렬 A 의 영공간 안에 있다면 $Ax = 0$ 이다. 이 식의 양변 왼쪽에 A^T 를 곱하면 $A^T Ax = 0$ 이 된다. 따라서 x 는 $A^T A$ 의 영공간 안에 있다.

이제 $A^T A$ 의 영공간에서 시작해보자. $A^T Ax = 0$ 에서 시작해서 $Ax = 0$ 임을 증명해야 한다. 그러나 일반적으로 존재한다고 확답할 수 없는 $(A^T)^{-1}$ 를 식 $A^T Ax = 0$ 의 양변 왼쪽에 곱할 수는 없다. 양변 왼쪽에는 다음과 같이 x^T 를 곱하기만 하면 된다.

$$(x^T)A^T Ax = 0 \quad \text{또는} \quad (Ax)^T(Ax) = 0 \quad \text{또는} \quad \|Ax\|^2 = 0 \quad (12)$$

위 식을 통해 $A^T Ax = 0$ 이면 Ax 의 길이가 0임을 보였다. 그러므로 $Ax = 0$ 이다. 따라서 행렬 A 의 영공간 안에 있는 모든 벡터 x 는 $A^T A$ 의 영공간 안에 있다. $A^T A$ 의 열이 선형종속인 경우, 행렬 A 의 열도 선형종속이다. 또한 $A^T A$ 의 열이 선형독립인 경우, 행렬 A 의 열도 선형독립이다. 이 경우 $A^T A$ 는 가역행렬이다. 따라서 다음이 성립한다.

행렬 A 의 열이 선형독립일 때, $A^T A$ 는 가역행렬, 대칭행렬, 정사각행렬이다. ■

다시 강조하면, $A^T A$ 의 크기는 $(n \times m) \times (m \times n)$ 이다. 그러면 $A^T A$ 는 $n \times n$ 정사각행렬이다. 또한 $(A^T A)^T = A^T (A^T)^T = A^T A$ 이므로, $A^T A$ 는 대칭행렬이다. 조금 전 행렬 A 의 열이 선형독립이면 $A^T A$ 가 가역행렬임을 증명했다. 다음 예를 통해 선형종속인 열과 선형독립인 열의 차이를 살펴보자.

$$\begin{array}{ccc} A^T & A & A^T A \\ \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix} & = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix} \\ \text{선형종속} & \text{특이행렬} & \end{array}$$

$$\begin{array}{ccc} A^T & A & A^T A \\ \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} & = \begin{bmatrix} 2 & 4 \\ 4 & 9 \end{bmatrix} \\ \text{선형독립} & \text{가역행렬} & \end{array}$$

지금까지 살펴본 내용을 간단히 요약해보자. 사영 $p = \hat{x}_1 a_1 + \dots + \hat{x}_n a_n$ 을 구하려면 $A^T A \hat{x} = A^T b$ 를 풀어야 한다. $\hat{x}$ 을 구하면 사영은 $p = A\hat{x}$ 이고 오차벡터는 $e = b - p = b - A\hat{x}$ 이다. 사영행렬 $P = A(A^T A)^{-1} A^T$ 를 알면 사영 $p = Pb$ 를 구할 수 있다. 사영행렬 P 는 $P = I$ 일 때만 가역행렬이다. 사영행렬 P 는 $P^2 = P$ 를 만족한다. 벡터 b 에서 부분공간 $C(A)$ 까지의 거리는 $\|e\|$ 이다.

핵심 아이디어 리뷰

- ① 벡터 b 의 벡터 a 방향 직선 위로의 사영은 $p = a\hat{x} = a \frac{a^T b}{a^T a}$ 이다.
- ② 랭크가 1인 사영행렬 $P = \frac{aa^T}{a^T a}$ 와 벡터 b 를 곱하면 사영 p 가 생성된다.
- ③ 벡터 b 를 부분공간 위로 사영하면 부분공간과 직교하는 $e = b - p$ 를 얻는다.
- ④ $m \times n$ 행렬 A 가 최대 랭크 n 을 가질 때, $A^T A \hat{x} = A^T b$ 를 풀면 $\hat{x}$ 과 $p = A\hat{x}$ 을 얻는다.
- ⑤ 사영행렬 $P = A(A^T A)^{-1} A^T$ 에 대해 $P^T = P$, $P^2 = P$, $Pb = p$ 가 성립한다.

4.2 확인문제

4.2A $b = (3, 4, 4)$ 를 $a = (2, 2, 1)$ 방향의 직선 위로 사영하고, $a^* = (1, 0, 0)$ 을 포함하는 평면 위로 사영하라. 첫 번째 오차벡터 $b - p$ 가 a 와 직교하는지 확인하고, 두 번째 오차벡터 $e^* = b - p^*$ 가 a^* 와 직교하는지 확인하라. a 와 a^* 를 포함하는 평면 위로의 사영에 대한 3×3 사영행렬 P 를 구하라. 그리고 이 평면 위로의 사영이 $p = 0$ 이 되는 벡터를 구하라.

풀이

$b = (3, 4, 4)$ 를 $a = (2, 2, 1)$ 방향의 직선 위로 사영하면 $p = 2a$ 이다.

$$\text{직선 위로의 사영} \quad p = \frac{a^T b}{a^T a} a = \frac{18}{9} (2, 2, 1) = (4, 4, 2) = 2a$$

오차벡터 $e = b - p = (-1, 0, 2)$ 는 $a = (2, 2, 1)$ 과 직교한다. 따라서 사영 p 를 정확히 구했음을 알 수 있다.

$a = (2, 2, 1)$ 과 $a^* = (1, 0, 0)$ 을 포함하는 평면은 행렬 $A = [a \ a^*]$ 의 열공간이다.

$$A = \begin{bmatrix} 2 & 1 \\ 2 & 0 \\ 1 & 0 \end{bmatrix}, \quad A^T A = \begin{bmatrix} 9 & 2 \\ 2 & 1 \end{bmatrix}, \quad (A^T A)^{-1} = \frac{1}{5} \begin{bmatrix} 1 & -2 \\ -2 & 9 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.8 & 0.4 \\ 0 & 0.4 & 0.2 \end{bmatrix}$$

이제 $p^* = Pb = (3, 4.8, 2.4)$ 이다. 오차벡터 $e^* = b - p^* = (0, -0.8, 1.6)$ 은 a, a^* 와 직교한다. 오차벡터 e^* 는 P 의 영공간 안에 있다. 따라서 e^* 의 사영은 0 이다. 이때 $P^2 = P = P^T$ 임에 주목하자.

4.2B 맥박이 분당 70회로 측정된 후, $x = 80, x = 120$ 으로 측정된다고 가정하자. 이 결과를 표현하는 세 방정식은 $Ax = b$ 이다. 여기서 x 는 하나의 미지수이고, $A^T = [1 \ 1 \ 1], b = (70, 80, 120)$ 이다. 가장 정확한 맥박 $\hat{x}$ 은 70, 80, 120의 _____이다. 미적분과 사영을 이용해 다음을 유도해보자.

- ① $\frac{dE}{dx} = 0$ 을 풀어서 $E = (x-70)^2 + (x-80)^2 + (x-120)^2$ 을 최소화하는 x 를 구한다.
- ② $b = (70, 80, 120)$ 을 $a = (1, 1, 1)$ 위로 사영하여 평균 $\hat{x} = 90$ 을 구한다.

4.2C $b_1, b_2, \dots, b_{999}$ 의 평균 $\hat{x}_{old}$ 를 알고 있다고 가정하자. b_{1000} 이 추가되면, 다음과 같이 새로운 평균 $\hat{x}_{new}$ 는 $\hat{x}_{old}$ 와 불일치 $b_{1000} - \hat{x}_{old}$ 의 선형결합이다.

$$\hat{x}_{new} = \frac{b_1 + \dots + b_{1000}}{1000} = \frac{b_1 + \dots + b_{999}}{999} + \frac{1}{1000} \left(b_{1000} - \frac{b_1 + \dots + b_{999}}{999} \right)$$

위 식은 이득행렬(gain matrix) $\frac{1}{1000}$ 을 갖는 '칼만 필터(Kalman filter)'로 다음과 같다.

$$\hat{x}_{new} = \hat{x}_{old} + \frac{1}{1000} (b_{1000} - \hat{x}_{old})$$

4.2 연습문제

0. 사영행렬 P 는 $P^2 = P$ 를 만족한다. 덧붙여 대칭 조건 $P^T = P$ 를 만족함에 따라 이 경우의 사영을 '직교사영(orthogonal projection)'이라고도 한다. 다음 물음을 통해 직교사영의 의미를 이해해보자.

- (a) $P = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, I - P = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$ 일 때, $P^2 = P$ 이고 $(I - P)^2 = I - P$ 임을 확인하라.
- (b) P 와 $I - P$ 의 열공간 안에 있는 벡터 v, w 를 찾으라. 두 열공간은 서로 직교하는가?
- (c) $P^2 = P, P^T = P$ 이면 Pv 와 $(I - P)w$ 의 내적이 0 이 됨을 보여라. 즉 $v^T P^T (I - P) w = 0$ 임을 보여라. 따라서 사영은 직교한다.

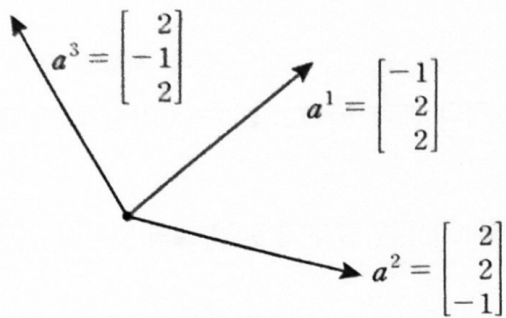
요약

사영 $p = Pb$ 가 오차벡터 $e = b - p$ 와 직교하려면 $P^2 = P = P^T$ 여야 한다.

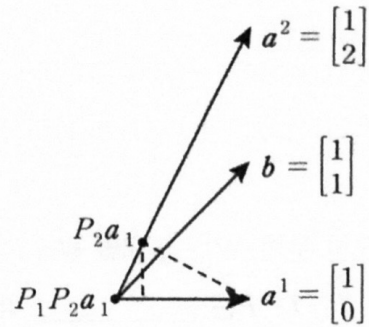
※ [문제 1~9]는 직선 위로의 사영 p , 오차벡터 $e = b - p$, 사영행렬 P 에 관한 문제이다.

1. 벡터 b 를 벡터 a 방향의 직선 위로 사영하라. 그리고 오차벡터 e 가 a 와 직교하는지 확인하라.
 (a) $b = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ (b) $b = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, a = \begin{bmatrix} -1 \\ -3 \\ -1 \end{bmatrix}$
2. 벡터 b 의 벡터 a 위로의 사영을 그림으로 나타내고, $p = \hat{x}a$ 형태로 표현하라.
 (a) $b = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (b) $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, a = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$
3. [문제 1]에서 각 벡터 a 방향 직선 위로의 사영행렬 $P = \frac{aa^T}{a^T a}$ 를 구하라. 두 경우 모두에 대해 $P^2 = P$ 임을 확인하라. 그리고 각 경우에 대해 Pb 를 계산하여 사영 p 를 구하라.
4. [문제 2]에서 각 벡터 a 방향 직선 위로의 사영행렬 P_1, P_2 를 구하라. 그리고 $(P_1 + P_2)^2 = P_1 + P_2$ 인지 확인하라. 이 식은 $P_1 P_2 = 0$ 일 때 성립할 것이다.
5. 두 벡터 $a_1 = (-1, 2, 2), a_2 = (2, 2, -1)$ 방향 직선 위로의 사영행렬 P_1, P_2 를 각각 구하라. 두 사영행렬 P_1, P_2 를 곱하고, 두 행렬의 곱이 무엇을 의미하는지 설명하라.
6. 벡터 $b = (1, 0, 0)$ 을 [문제 5]의 벡터 a_1, a_2 방향 직선 위로 사영한 p_1, p_2 를 각각 구하라. 또한 벡터 $b = (1, 0, 0)$ 을 벡터 $a_3 = (2, -1, 2)$ 위로 사영한 p_3 도 구하라. 세 사영의 합 $p_1 + p_2 + p_3$ 를 구하라.
7. [문제 5]~[문제 6]의 과정을 계속해서, $a_3 = (2, -1, 2)$ 위로 사영하는 사영행렬 P_3 를 구하라. 그리고 $P_1 + P_2 + P_3 = I$ 임을 확인하라. 이 식은 기저벡터 a_1, a_2, a_3 가 서로 직교하기 때문에 성립한다.
8. 벡터 $b = (1, 1)$ 을 벡터 $a_1 = (1, 0), a_2 = (1, 2)$ 방향 직선 위로 사영한 p_1, p_2 를 각각 구하라. 사영 p_1, p_2 를 그림으로 나타내고 합 $p_1 + p_2$ 를 구하라. 벡터 a_1, a_2 는 서로 직교하지 않으므로, $p_1 + p_2$ 는 벡터 b 가 되지 않는다.
9. [문제 8]에서 벡터 b 의 a_1, a_2 가 속한 평면 위로의 사영은 b 이다. 가역행렬 $A = [a_1 \ a_2] = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ 에 대한 사영행렬 $P = A(A^T A)^{-1} A^T$ 를 구하라.

10. 벡터 $a_1 = (1, 0)$ 을 벡터 $a_2 = (1, 2)$ 위로 사영하라. 그리고 사영 결과를 다시 벡터 a_1 위로 사영하라. 두 사영을 그림으로 나타내고, 사영행렬의 곱 $P_1 P_2$ 를 구하라. 이렇게 곱한 결과도 사영이 되는가?



[그림 4.4] [문제 5~7] 직교하는 a_1, a_2, a_3



[그림 4.5] [문제 8~10] 직교하지 않는 a_1, a_2, a_3

※ [문제 11~21]은 부분공간 위로의 사영과 사영행렬에 관한 문제이다.

11. $m \times n$ 행렬 P 가 $\mathbb{R}^m$ 안의 벡터를 부분공간 S 위로 사영한다고 할 때, 행렬 P 의 4가지 기본 부분공간을 구하라.
12. $A^T A \hat{x} = A^T b$ 를 풀고 $p = A \hat{x}$ 을 계산하여 벡터 b 를 행렬 A 의 열공간 위로 사영하라. 그리고 $e = b - p$ 를 계산하라. 오차벡터 e 는 행렬 A 의 열과 직교해야 한다.
- (a) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, $b = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$, $b = \begin{bmatrix} 4 \\ 4 \\ 6 \end{bmatrix}$
13. [문제 12]의 열공간 위로의 사영행렬 P_1, P_2 를 구하라. 그리고 $P_1 b$ 가 사영 p_1 이 됨을 확인하라. 또한 $P_2^2 = P_2$ 임을 확인하라.
14. (꼭 풀어볼 만한 문제) A 를 4×4 항등행렬에서 마지막 열을 제거한 행렬이라 가정하자. 따라서 A 는 4×3 행렬이다. 벡터 $b = (1, 2, 3, 4)$ 를 행렬 A 의 열공간 위로 사영하라. 사영행렬 P 는 어떤 형태이며 무엇인가?
15. 벡터 b 는 행렬 A 의 첫 번째 열에 2를 곱한 것이다. 벡터 b 의 행렬 A 의 열공간 위로의 사영을 구하라. 이 경우 $P = I$ 인가? $b = (0, 2, 4)$ 이고 행렬 A 의 열이 $(0, 1, 2), (1, 2, 0)$ 일 때, 사영 p 와 사영행렬 P 를 구하라.

16. 행렬 A 가 2배가 되면 사영행렬은 $P = 2A(4A^T A)^{-1}2A^T$ 가 된다. 이는 $A(A^T A)^{-1}A^T$ 와 같다. 행렬 $2A$ 의 열공간과 행렬 A 의 열공간 사이에 어떤 관계가 성립하는가? $\hat{x}$ 은 행렬 A 와 $2A$ 에 대해 동일한가?
17. 두 벡터 $(1, 2, -1)$, $(1, 0, 1)$ 의 선형결합들 중 어떤 것이 벡터 $b = (2, 1, 1)$ 과 가장 가까운가?
18. (중요한 문제) $P^2 = P$ 이면 $(I - P)^2 = I - P$ 임을 보여라. 사영행렬 P 가 행렬 A 의 열공간 위로 사영할 때, $I - P$ 는 _____ 위로 사영한다.
19. 다음 물음에 답하라.
- (a) P 가 벡터 $(1, 1)$ 방향의 직선 위로 사영하는 2×2 행렬일 때,
 $I - P$ 는 _____ 위로 사영하는 행렬이다.
- (b) P 가 벡터 $(1, 1, 1)$ 방향의 직선 위로 사영하는 3×3 행렬일 때,
 $I - P$ 는 _____ 위로 사영하는 행렬이다.
20. 평면 $x - y - 2z = 0$ 위로의 사영행렬을 찾기 위해, 해당 평면 위에서 두 벡터를 선택하여 행렬 A 의 열로 만들어라. 그러면 해당 평면은 행렬 A 의 열공간이 될 것이다. $P = A(A^T A)^{-1}A^T$ 를 구하라.
21. [문제 20]과 동일한 평면 $x - y - 2z = 0$ 위로의 사영행렬 P 를 찾기 위해, 해당 평면과 직교하는 벡터 e 를 찾으라. 그리고 사영행렬 $Q = \frac{ee^T}{e^T e}$ 와 $P = I - Q$ 를 구하라.

※ [문제 22~27]에서는 사영행렬 P 가 $P^2 = P$ 와 $P^T = P$ 를 만족함을 보인다.

22. 행렬 $P = A(A^T A)^{-1}A^T$ 를 자기 자신과 곱하고(P^2), $P^2 = P$ 임을 증명하라. Pb 의 사영인 $P(Pb)$ 가 Pb 와 항상 같은 이유를 설명하라.
23. P^T 를 계산하여 $P = A(A^T A)^{-1}A^T$ 가 대칭행렬임을 증명하라. 대칭행렬의 역행렬은 대칭행렬임을 기억하라.
24. A 가 가역인 정사각행렬이면, [EX 3]의 바로 뒤에서 $(A^T A)^{-1}$ 의 분할에 대해 설명한 주의사항이 적용되지 않는다. 즉 $AA^{-1}(A^T)^{-1}A^T = I$ 가 성립한다. A 가 가역행렬일 때, 왜 $P = I$ 가 성립하는가? 이때 오차벡터 e 는 무엇인가?

25. A^T 의 영공간은 열공간 $C(A)$ 와 어떤 관계인가? 만약 $A^T b = 0$ 이면, 벡터 b 의 $C(A)$ 위로의 사영은 $p = \underline{\hspace{2cm}}$ 이어야 한다. $A(A^T A)^{-1} A^T b = p$ 임을 확인하라.
26. $\mathbb{R}^m$ 의 n 차원 부분공간 S 위로의 사영행렬 P 의 랭크는 $r = n$ 이다. 그 이유는 사영 Pb 가 부분공간 S 를 채우기 때문이다. 그러므로 S 는 P 의 $\underline{\hspace{2cm}}$ 이다.
27. $m \times m$ 행렬 A 가 $A^2 = A$ 를 만족하고 랭크가 m 일 때, $A = I$ 임을 어떻게 알 수 있는가?
28. 이 절에서 살펴본 내용 중 알아두어야 할 중요한 사실로 ' $A^T Ax = 0$ 이면 $Ax = 0$ '이 있다. 이 명제를 새로 증명해보자. 벡터 Ax 는 $\underline{\hspace{2cm}}$ 의 영공간 안에 있다. 그리고 벡터 Ax 는 항상 $\underline{\hspace{2cm}}$ 의 열공간 안에 있다. 직교하는 두 공간 모두 안에 속하려면 Ax 는 0 이어야 한다.
29. $P^T = P$, $P^2 = P$ 를 이용해, 사영행렬 P 의 2열의 크기의 제곱이 항상 주대각성분 P_{22} 와 같음을 증명하라. 또한 이를 만족하는 예를 제시하라.
30. 행렬 B 의 최대 행 랭크가 m 이고 선형독립인 행을 가질 때, BB^T 가 가역행렬임을 보여라.

도전문제

31. 다음 물음에 답하라.
- (a) 행렬 $A = \begin{bmatrix} 3 & 6 & 6 \\ 4 & 8 & 8 \end{bmatrix}$ 의 열공간 위로의 사영행렬 P_C 를 구하라.
- (b) (a)의 행렬 A 의 행공간 위로의 사영에 대한 3×3 사영행렬 P_R 을 구하라. 그리고 $B = P_C A P_R$ 을 계산하라. 또한 행렬 B 가 어떤 특징을 갖는지 설명하라.
32. $\mathbb{R}^m$ 에서 벡터 b 와 $a_1, \dots, a_n$ 의 선형결합 p 가 주어진다고 하자. 벡터 b 를 $a_1, \dots, a_n$ 으로 생성되는 부분공간 위로 사영한 결과가 p 임을 어떻게 판정할 수 있는가?
33. P_1 은 다음과 같은 행렬 A 의 1열에 의해 생성되는 1차원 부분공간 위로의 사영행렬이다. P_2 는 행렬 A 의 2차원 열공간 위로의 사영행렬이다. $P_2 P_1$ 을 계산하라.

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$$

34. P_1, P_2 는 사영행렬이다($P_i^2 = P_i = P_i^T$). $P_1 P_2$ 가 사영행렬일 필요충분조건은 $P_1 P_2 = P_2 P_1$ 이 성립하는 것임을 증명하라.

4.3 최소제곱근사

1. $A^T A \hat{x} = A^T b$ 를 풀면, 벡터 b 의 행렬 A 의 열공간 위로의 사영 $p = A \hat{x}$ 을 구할 수 있다.
2. $Ax = b$ 의 해가 존재하지 않을 때, $\hat{x}$ 은 '최소제곱해'이다. 이때 $\|A\hat{x} - b\|^2 = (\text{최솟값})$ 이다.
3. $E = \|Ax - b\|^2$ 의 도함수를 0이라 하면 $\left(\frac{\partial E}{\partial x_i} = 0\right)$, $A^T A \hat{x} = A^T b$ 가 유도된다.
4. 점 $(t_1, b_1), \dots, (t_m, b_m)$ 을 하나의 직선에 맞추기 위해, 행렬 A 는 $(1, \dots, 1), \dots, (t_1, \dots, t_m)$ 을 열로 갖는다.
5. 위와 같은 경우 $A^T A$ 는 2×2 행렬 $\begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix}$ 이고, $A^T b$ 는 벡터 $\begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix}$ 이다.

종종 $Ax = b$ 의 해가 존재하지 않는다. 이에 대한 일반적인 이유는 너무 많은 방정식이 존재하기 때문이다. 이는 행렬 A 에 열보다 행이 더 많은 경우이다. 즉 미지수보다 방정식이 더 많은 경우이다($m > n$). n 개의 열들은 m 차원 공간의 작은 부분공간을 생성한다. 모든 측정값이 완벽하지 않다면, 벡터 b 는 행렬 A 의 열공간 밖에 있다. 소거 과정은 해가 존재하지 않는 방정식에 도달하면 멈춘다. 그러나 측정값에 잡음(noise)이 포함된 경우, 단순히 멈출 수는 없다.

반복해서 설명하자면, 오차벡터 $e = b - Ax$ 를 항상 0으로 줄일 수는 없다. e 가 0일 때 x 는 $Ax = b$ 의 정확한 해이다. e 의 크기가 가능한 한 작아서 $\|b - A\hat{x}\|^2$ 이 최솟값에 도달하면, $\hat{x}$ 은 최소제곱해가 된다. 이 절의 목표는 $\hat{x}$ 을 계산하고 사용하는 것이다. 이와 같은 문제들은 실제 문제이므로 반드시 해답이 필요하다.

4.2절에서는 사영 p 를 강조했다. 이 절에서는 최소제곱해 $\hat{x}$ 을 강조해 설명한다. p 와 $\hat{x}$ 은 $p = A\hat{x}$ 으로 연결된다. 기본 방정식은 여전히 정규방정식 $A^T A \hat{x} = A^T b$ 이다. 다음은 '정규방정식'에 도달하는 비공식적인 방법이다.

$Ax = b$ 의 해가 존재하지 않으면, 이 식의 양변에 A^T 를 곱하고 $A^T A \hat{x} = A^T b$ 를 푼다.

EX 1 최소제곱과 관련한 중요한 응용은 m 개의 점을 직선에 맞추는 문제이다. 우선 세 점으로 시작한다. 세 점 $(0, 6)$, $(1, 0)$, $(2, 0)$ 과 가장 가까운 직선을 찾아보자.

주어진 세 점을 지나는 직선 $b = C + Dt$ 는 존재하지 않는다. 이 문제에서는 세 방정식을 만족하는 두 수 C , D 가 필요하다(미지수가 $n = 2$ 개, 방정식이 $m = 3$ 개인 경우). 다음은 주어진 값 $b = 6, 0, 0$ 에 맞추기 위한 $t = 0, 1, 2$ 에 대한 세 방정식이다.

$t = 0$ 인 경우, $C + D \cdot 0 = 6$ 이면 첫 번째 점은 직선 $b = C + Dt$ 위에 있다.

$t = 1$ 인 경우, $C + D \cdot 1 = 0$ 이면 두 번째 점은 직선 $b = C + Dt$ 위에 있다.

$t = 2$ 인 경우, $C + D \cdot 2 = 0$ 이면 세 번째 점은 직선 $b = C + Dt$ 위에 있다.

이와 같은 3×2 시스템(연립방정식)의 해는 존재하지 않는다. 즉 벡터 $b = (6, 0, 0)$ 은 열 $(1, 1, 1)$ 과 $(0, 1, 2)$ 의 선형결합이 아니다. 위와 같은 연립방정식으로부터 A , x , b 의 값을 도출하면 다음과 같다. 방정식 $Ax = b$ 의 해는 존재하지 않는다.

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad x = \begin{bmatrix} C \\ D \end{bmatrix}, \quad b = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

4.2절의 [EX 3]에서는 위와 동일한 행렬과 벡터를 살펴봤고, $\hat{x} = (5, -3)$ 을 구했다. 이 수가 바로 최적의 C , D 이다. 따라서 $5 - 3t$ 가 세 점에 가장 적합한 직선이 될 것이다. 한편 $A^T A \hat{x} = A^T b$ 가 성립하는 이유를 설명함으로써, 사영을 최소제곱과 연결해야 한다.

실제 응용 문제에서는 $m = 3$ 이 아니라 $m = 100$ 개의 점이 주어지는 경우가 흔하다. 이 점들은 어떤 직선 $C + Dt$ 에도 정확히 맞출 수 없다. 앞서 사용한 수 $6, 0, 0$ 은 오차를 강조해서 보여주기 위한 값으로, [그림 4.6]에서 오차 e_1, e_2, e_3 를 확인할 수 있다.

오차 최소화

오차 $e = b - Ax$ 를 가능한 한 작게 만들려면 어떻게 해야 할까? 이는 굉장히 중요한 질문이며, 아름다운 해법이 존재한다. 최적의 x (이를 $\hat{x}$ 이라 함)는 기하학적으로 구할 수 있다(오차가 행렬 A 의 열공간과 직교하기 때문). 대수학적으로는 $A^T A \hat{x} = A^T b$ 에서 그 해답을 얻을 수 있다. 또한 미분적분학을 통해서도 동일한 결과 $\hat{x}$ 를 얻을 수 있다(오차 $\|Ax - b\|^2$ 의 도함수가 $\hat{x}$ 에서 0이 되기 때문).

먼저 기하학적 관점에서 [EX 1]을 예로 생각해보자. 모든 Ax 는 열 $(1, 1, 1)$ 과 $(0, 1, 2)$ 가 포함

된 평면 위에 있다. 이 평면에서 벡터 b 와 가장 가까운 점을 찾아보자. 가장 가까운 점은 사영 p 이다.

$A\hat{x}$ 에 대한 최적의 선택은 p 이다. 가능한 한 가장 작은 오차는 $e = b - p$ 이고, 이는 행렬 A 의 열과 직교한다. p 는 행렬 A 의 열공간 안에 있으므로, $p = (p_1, p_2, p_3)$ 에서 세 점은 하나의 직선 위에 있다. 직선에 맞추기 위한 최적의 선택은 $\hat{x} = (C, D)$ 이다.

다음으로 대수학적 관점에서 생각해보자. 모든 벡터 b 는 두 부분으로 분할된다. 열공간에서의 부분은 p 이고, 수직 부분은 e 이다. 한편, 다음과 같이 풀 수 없는 방정식 $Ax = b$ 와 (e 를 제거하고 $A^T A \hat{x} = A^T b$ 를 풀어냄으로써) 풀 수 있는 방정식 $A\hat{x} = p$ 가 있다.

$$Ax = b = p + e : \text{풀 수 없음}, \quad A\hat{x} = p : \text{풀 수 있음}, \quad \hat{x} = (A^T A)^{-1} A^T b \quad (1)$$

$A\hat{x} = p$ 의 해는 다음과 같이 최소의 오차 e 를 남긴다.

$$\text{임의의 } x \text{에 대한 제곱오차} \quad \|Ax - b\|^2 = \|Ax - p\|^2 + \|e\|^2 \quad (2)$$

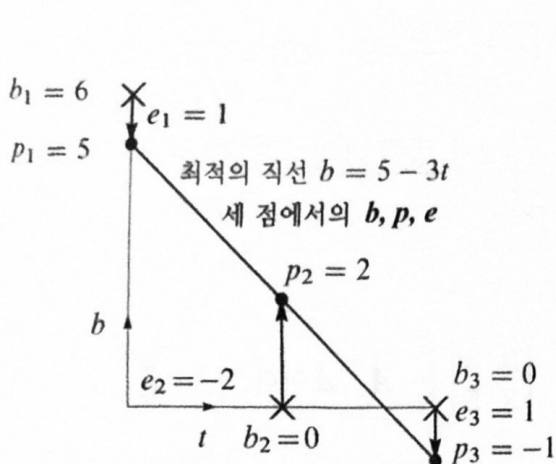
식 (2)는 직각삼각형에 대한 피타고라스 정리 $c^2 = a^2 + b^2$ 과 같다. 열공간의 벡터 $Ax - p$ 는 좌영공간의 벡터 e 와 직교한다. $x = \hat{x}$ 으로 선택하면 $Ax - p$ 를 0으로 줄일 수 있다. 이에 따라 식 (2)에는 더 이상 줄일 수 없는 최소의 오차 $e = (e_1, e_2, e_3)$ 만 남는다.

‘최소’의 의미에 주목하자. $Ax - b$ 의 제곱크기는 다음과 같이 최소화된다.

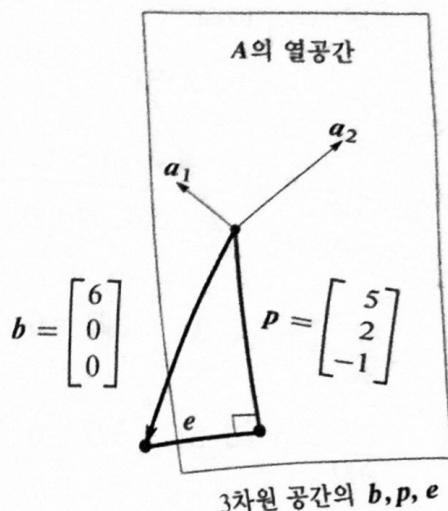
$$\text{최소제곱해 } \hat{x} \text{은 } E = \|Ax - b\|^2 \text{을 가능한 한 작게 만든다.}$$

[그림 4.6(a)]는 세 점과 가장 가까운 직선(최소제곱선)을 보여준다. b_i 와 p_i 의 거리 차이 $e_1, e_2, e_3 = 1, -2, 1$ 에 따라 이 직선은 주어진 세 점을 못 지난다. 이때의 거리는 수직 거리이다. 이와 같은 최소제곱선은 $E = e_1^2 + e_2^2 + e_3^2$ 을 최소화한다.

[그림 4.6(b)]는 [그림 4.6(a)]와 동일한 문제를 3차원 공간(bpe -공간)에서 보여준다. 벡터 b 는 행렬 A 의 열공간 안에 없다. 그래서 $Ax = b$ 를 풀 수 없었다. 세 점을 지나는 직선은 없다. 가능한 한 가장 작은 오차는 수직인 벡터 e 이다. 이는 $e = b - A\hat{x}$ 으로, 세 방정식에 대한 오차벡터 $(1, -2, 1)$ 이다. $(1, -2, 1)$ 은 최적의 직선으로부터의 거리이다. 다음 두 그림에 대한 배경에는 정규방정식 $A^T A \hat{x} = A^T b$ 가 있다.



(a) 오차 = (직선까지의 수직 거리)



(b) 오차 $e = (1, -2, 1) = b - p$

[그림 4.6] 최적의 직선과 사영(동일한 문제에 대한 두 그림)

[그림 4.6]을 보면 직선에 대한 높이는 $p = (5, 2, -1)$ 이고 오차는 $e = (1, -2, 1)$ 이다. 방정식 $A^T A \hat{x} = A^T b$ 를 풀면 $\hat{x} = (5, -3)$ 이 나오는데, 이 결과는 [EX 1]과 동일하다. 따라서 최적의 직선은 $b = 5 - 3t$ 이고, 가장 가까운 점은 $p = 5a_1 - 3a_2$ 이다.

오차벡터의 성분 1, -2, 1을 모두 더하면 0이 된다. 그 이유는 오차벡터 $e = (e_1, e_2, e_3)$ 는 행렬 A 의 1열 $(1, 1, 1)$ 과 서로 직교하기 때문이다. 즉 두 벡터 e 와 $(1, 1, 1)$ 의 내적은 $e_1 + e_2 + e_3 = 0$ 이다.

마지막으로 미분적분학의 관점에서 생각해보자. 대부분의 함수에 대한 최소화 문제는 미분적분학을 통해 해결한다. 최소화되는 지점에서 그래프는 최저점에 이르고, 모든 방향에서의 도함수는 0이 된다. 여기서 최소화해야 할 오차 E 는 다음과 같이 제곱의 합 $e_1^2 + e_2^2 + e_3^2$ (세 방정식의 오차제곱)이다.

$$E = \|Ax - b\|^2 = (C + D \cdot 0 - 6)^2 + (C + D \cdot 1)^2 + (C + D \cdot 2)^2 \quad (3)$$

식 (3)에서 미지수는 C 와 D 이다. 두 미지수를 가진 두 도함수가 존재하며, 두 도함수 모두 최소점에서 0이다. $\frac{\partial E}{\partial C}$ 에서는 D 를 상수로 취급하고, $\frac{\partial E}{\partial D}$ 에서는 C 를 상수로 취급하므로 이 도함수들은 '편도함수'이다. 다음은 식 (3)의 E 에 대한 도함수이다.

$$\frac{\partial E}{\partial C} = 2(C + D \cdot 0 - 6) + 2(C + D \cdot 1) + 2(C + D \cdot 2) = 0$$

$$\frac{\partial E}{\partial D} = 2(C + D \cdot 0 - 6)(0) + 2(C + D \cdot 1)(1) + 2(C + D \cdot 2)(2) = 0$$

$\frac{\partial E}{\partial D}$ 의 식에는 연쇄법칙에 따라 추가 인수 0, 1, 2가 포함된다(앞 식을 보면 $(C+2D)^2$ 의 마지막 도함수는 $2(C+D \cdot 2)(2)$ 이다). 이러한 추가 인수가 $\frac{\partial E}{\partial C}$ 에는 포함되지 않는다.

$\|Ax - b\|^2$ 의 도함수에 있는 추가 인수 1, 1, 1과 0, 1, 2가 행렬 A 의 열인 것은 우연이 아니다. 이제 앞의 도함수에 대한 식을 정리하면 다음과 같다.

$$\begin{aligned} \frac{\partial E}{\partial C} = 0 &\Leftrightarrow 3C + 3D = 6 \\ \frac{\partial E}{\partial D} = 0 &\Leftrightarrow 3C + 5D = 0 \end{aligned} \quad \text{행렬 } \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \text{는 } A^T A \text{이다.} \quad (4)$$

식 (4)의 두 방정식은 $A^T A \hat{x} = A^T b$ 이다. 최적의 C, D 는 $\hat{x}$ 의 성분이다. 즉 미분적분학의 방정식은 선형대수학의 '정규방정식'과 동일하다. 다음은 $\|Ax - b\|^2 = x^T A^T A x - 2x^T A^T b + b^T b$ 를 최소화하기 위한 방정식이다.

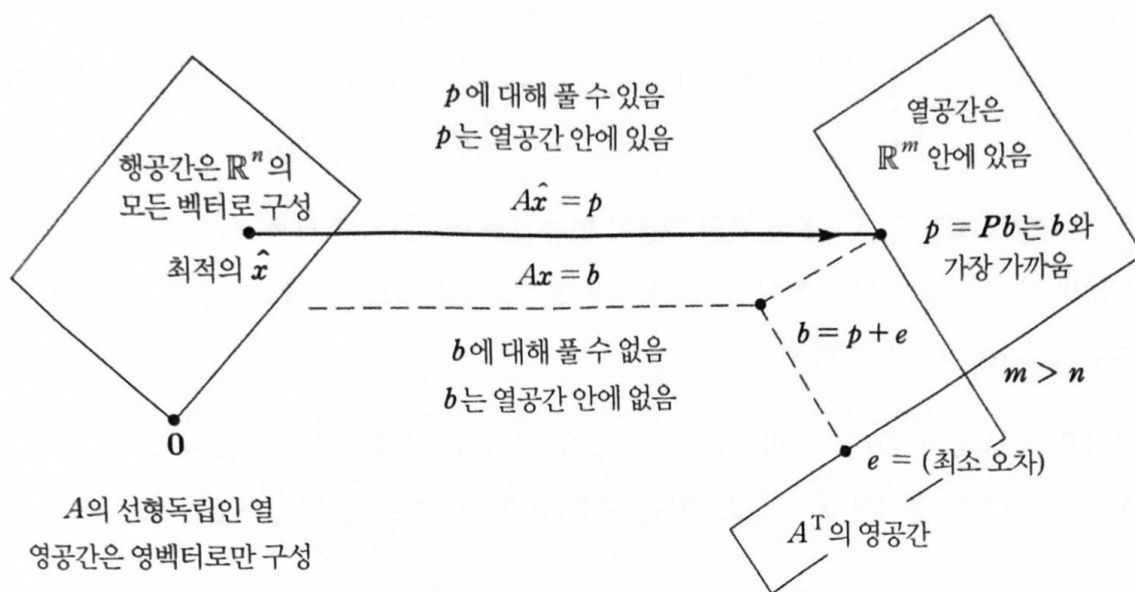
$$A^T A \hat{x} = A^T b \text{ 일 때, } \|Ax - b\|^2 \text{의 편도함수는 0이다.}$$

방정식의 해는 $C = 5, D = -3$ 이다. 따라서 $b = 5 - 3t$ 가 세 점과 가장 가까운 최적의 직선이다. $t = 0, 1, 2$ 일 때, 이 직선은 $p = (5, 2, -1)$ 을 지나고, $b = (6, 0, 0)$ 을 지나지 않는다. 오차는 $(1, -2, 1)$ 이고, 이것이 바로 벡터 e 이다.

최소제곱에 대한 큰 그림

이 책의 주요 그림인 [그림 4.1]은 4가지 기본 부분공간과 행렬 A 의 실제 작용을 보여준다. [그림 4.1]의 왼쪽에 있는 벡터 x 는 오른쪽에 있는 $b = Ax$ 로 이동했다. 또한 모든 벡터 x 는 $x_r + x_n$ 으로 분할되었다. $Ax = b$ 의 해는 많이 존재했다.

이 절에서는 상황이 정반대이다. $Ax = b$ 의 해는 존재하지 않는다. 여기서는 벡터 x 를 분할하는 대신에 $b = p + e$ 를 분할한다. [그림 4.7]은 최소제곱에 대한 큰 그림을 보여준다. 여기서는 $Ax = b$ 대신에 $A\hat{x} = p$ 를 푼다. 오차 $e = b - p$ 는 피할 수 없다.



[그림 4.7] 사영 $p = A\hat{x}$ 이 벡터 b 와 가장 가까우므로, $\hat{x}$ 은 $E = \|b - Ax\|^2$ 을 최소화한다.

영공간 $N(A)$ 가 얼마나 작은지에 주목하라. $N(A)$ 는 단 하나의 점만 포함한다. 선형독립인 열을 이용하면 $Ax = 0$ 의 유일한 해는 $x = 0$ 임을 알 수 있다. 그러면 $A^T A$ 는 가역행렬이다. 방정식 $A^T A\hat{x} = A^T b$ 는 최적의 벡터 $\hat{x}$ 을 완전히 결정한다. 오차 $e = b - p$ 는 $A^T e = 0$ 을 만족한다.

모든 벡터 x 는 $x_r + x_n$ 으로 분할되고, 모든 벡터 b 는 $p + e$ 로 분할된다. 최적의 해는 행공간 안에 있는 $\hat{x} = \hat{x}_r$ 이다. 오차 e 는 피할 수 없고, 영공간 안에 있는 x_n 은 필요하지 않다. 결국 남는 것은 $A\hat{x} = p$ 이다.

직선 맞추기

직선 맞추기(fitting a line)는 최소제곱법에 대한 가장 명확한 응용이다. 보통 $m > 2$ 개의 점으로 시작하며, 이 점들이 하나의 직선 근처에 있기를 기대한다. 시간 $t_1, \dots, t_m$ 에서 각 m 개의 점의 높이는 $b_1, \dots, b_m$ 이다. 최적의 직선 $C + Dt$ 는 각 점과 수직 거리 $e_1, \dots, e_m$ 만큼 어긋나 있어 각 점들을 지나지 못한다. 완벽한 직선은 없으며, 최소제곱선은 $E = e_1^2 + \dots + e_m^2$ 을 최소화한다.

이 절의 [EX 1]에는 [그림 4.6]과 같이 세 점이 있었다. 이제 (큰 수인) m 개의 점을 가정해보자. $\hat{x}$ 의 두 성분은 여전히 C, D 이다.

$Ax = b$ 의 해가 존재할 때, 직선은 m 개의 점을 모두 지난다. 그러나 일반적으로 이는 불가능하다. 두 미지수 C, D 가 직선을 결정하므로, 행렬 A 는 오직 $n = 2$ 개의 열만 갖는다. 이제 m 개의 점을 최대한 직선에 맞추기 위해, m 개의 방정식을 풀고자 한다(미지수는 2개뿐이다).

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \text{ 일 때, } Ax = b \Leftrightarrow \begin{cases} C + Dt_1 = b_1 \\ C + Dt_2 = b_2 \\ \vdots \\ C + Dt_m = b_m \end{cases} \quad (5)$$

열공간이 너무 얇기 때문에 벡터 b 는 거의 확실히 열공간 밖에 있다. 벡터 b 가 열공간 안에 있으면 모든 점이 직선 위에 놓이게 된다. 이러한 경우 $b = p$ 이다. 그러면 $Ax = b$ 를 풀 수 있고, 이때 오차는 $e = (0, \dots, 0)$ 이다.

가장 가까운 직선 $C + Dt$ 는 높이 $p_1, \dots, p_m$ 을 가지며, 이때 오차는 $e_1, \dots, e_m$ 이다.

$A^T A \hat{x} = A^T b$ 를 풀어서 $\hat{x} = (C, D)$ 를 구한다. 각 오차는 $e_i = b_i - C - Dt_i$ 이다.

직선에 점들을 맞추는 것은 매우 중요하므로, 두 방정식 $A^T A \hat{x} = A^T b$ 를 한 번에 정리하여 제시하고자 한다. (모든 시간 t_i 가 동일하지 않다면) 행렬 A 의 두 열은 선형독립이다. 따라서 최소 제곱 문제로 바뀌서 $A^T A \hat{x} = A^T b$ 를 푼다.

$$\text{내적 행렬} \quad A^T A = \begin{bmatrix} 1 & \dots & 1 \\ t_1 & \dots & t_m \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix} \quad (6)$$

정규방정식의 우변에는 다음과 같이 2×1 벡터 $A^T b$ 가 있다.

$$A^T A \begin{bmatrix} C \\ D \end{bmatrix} = A^T A \hat{x} = A^T b = \begin{bmatrix} 1 & \dots & 1 \\ t_1 & \dots & t_m \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix} \quad (7)$$

실제 문제에서는 식 안의 미지수에 대한 수가 주어진다. 최적의 $\hat{x} = (C, D)$ 는 $(A^T A)^{-1} A^T b$ 이다. 직선 $C + Dt$ 는 다음과 같이 $A^T A \hat{x} = A^T b$ 일 때, $e_1^2 + \dots + e_m^2 = \|Ax - b\|^2$ 을 최소화한다.

$$A^T A \hat{x} = A^T b \Leftrightarrow \begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix} \quad (8)$$

직선 위의 m 개의 점에서 수직 오차는 $e = b - p$ 의 성분이다. 오차벡터 e (즉 잔차(residual) $b - A\hat{x}$)는 행렬 A 의 열과 직교한다(기하학 관점). 오차는 A^T 의 영공간 안에 있다(선형대수학 관점). 최적의 $\hat{x} = (C, D)$ 는 다음의 제곱합인 전체 오차 E 를 최소화한다(미분적분학 관점).

$$E(x) = \|Ax - b\|^2 = (C + Dt_1 - b_1)^2 + \dots + (C + Dt_m - b_m)^2$$

미분적분학에서는 도함수 $\frac{\partial E}{\partial C}, \frac{\partial E}{\partial D}$ 를 0으로 놓고 $A^T A \hat{x} = A^T b$ 를 유도한다.

다른 최소제곱 문제의 경우 2개 이상의 미지수를 갖는다. 최적의 포물선을 구하기 위해서는 $n = 3$ 개의 계수 C, D, E 가 필요하다([EX 2] 참고). 일반적으로 n 개의 매개변수 $x_1, \dots, x_n$ 에 따라 m 개의 데이터 점을 맞춘다. 행렬 A 에는 n 개의 열이 있고 $n < m$ 이다. $\|Ax - b\|^2$ 의 도함수를 구하면 n 개의 방정식 $A^T A \hat{x} = A^T b$ 를 얻는다. 제곱함수의 도함수는 선형이다. 이러한 이유로 최소제곱법이 널리 사용되고 있다.

EX 2 측정 시간 t_i 들의 합 0일 때, 행렬 A 는 직교하는 열을 갖는다. $t = -2, 0, 2$ 일 때 $b = 1, 2, 4$ 라 가정하자. 행렬 A 의 열들의 내적은 0이다(즉 $(1, 1, 1)$ 은 $(-2, 0, 2)$ 와 직교한다).

$$\begin{aligned} C + D(-2) &= 1 \\ C + D(0) &= 2 \\ C + D(2) &= 4 \end{aligned} \quad \text{또는} \quad Ax = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

행렬 A 의 열이 서로 직교할 때, 다음과 같이 $A^T A$ 는 대각행렬이 된다.

$$A^T A \hat{x} = A^T b \quad \Leftrightarrow \quad \begin{bmatrix} 3 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} \quad (9)$$

여기서 중요한 사실은, $A^T A$ 가 대각행렬이므로 $C = \frac{7}{3}$ 과 $D = \frac{6}{8}$ 을 따로 풀 수 있다는 점이다. $A^T A$ 안에 있는 0은 행렬 A 의 직교하는 열들의 내적이다. $m = 3$ 과 $t_1^2 + t_2^2 + t_3^2 = 8$ 을 만족하는 대각행렬 $A^T A$ 는 항등행렬만큼 간단하다.

직교하는 열은 매우 유용하므로, 측정 시간에서 평균 시간 $\hat{t} = \frac{t_1 + \dots + t_m}{m}$ 을 뺀으로써 시간을 이동하는 것이 좋다. 원래 시간이 1, 3, 5라면 이들의 평균은 $\hat{t} = 3$ 이다. 이동한 시간 $T = t - \hat{t} = t - 3$ 을 모두 더하면 0이 된다.

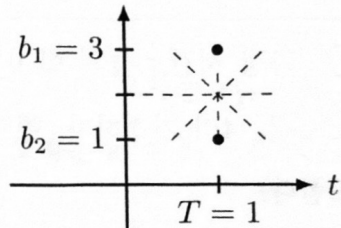
$$\begin{aligned} T_1 &= 1 - 3 = -2 \\ T_2 &= 3 - 3 = 0 \\ T_3 &= 5 - 3 = 2 \end{aligned} \quad A_{\text{new}} = \begin{bmatrix} 1 & T_1 \\ 1 & T_2 \\ 1 & T_3 \end{bmatrix} \quad A_{\text{new}}^T A_{\text{new}} = \begin{bmatrix} 3 & 0 \\ 0 & 8 \end{bmatrix}$$

이제 간단한 방정식 (9)로부터 C, D 를 얻는다. 그러면 최적의 직선은 $C + DT$, 즉 $C + D(t - \hat{t}) = C + D(t - 3)$ 이다. 이 절의 [연습문제 30]에서는 C, D 에 대한 공식을 제공한다. 그 공식은 다음 절에서 살펴볼 ‘그람-슈미트 아이디어’의 완벽한 예이다. 즉 이 방법은 미리 열들을 직교하게 만드는 것이다. 그러면 $A_{\text{new}}^T A_{\text{new}}$ 가 대각행렬이 되어 $\hat{x}_{\text{new}}$ 를 쉽게 구할 수 있다.

A의 열이 선형종속일 때 $\hat{x}$ 은 무엇인가

이 장에서는 처음부터 행렬 A 의 열들이 선형독립이라고 가정했다. 그러면 $A^T A$ 는 가역행렬이다. 이에 따라 $A^T A \hat{x} = A^T b$ 는 $Ax = b$ 의 최소제곱해 $\hat{x}$ 을 제공한다.

행렬 A 가 선형종속인 열을 갖는 경우 최적의 $\hat{x}$ 은 무엇인가? 다음 예를 통해 살펴보자.

$$\begin{aligned} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 3 \\ 1 \end{bmatrix} = b & \quad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = p \\ Ax &= b & \quad A\hat{x} = p \end{aligned}$$


위 그림을 보면 측정값 $b_1 = 3$ 과 $b_2 = 1$ 은 동일한 시간 T 에서 측정된다. 직선 $C + Dt$ 는 두 점을 모두 지날 수 없다. 따라서 $b = (3, 1)$ 을 행렬 A 의 열공간 안에 있는 $p = (2, 2)$ 로 사영해야 한다. 이를 통해 방정식 $Ax = b$ 를 $A\hat{x} = p$ 로 바꾼다. 이로써 해를 갖지 않는 방정식이 무수히 많은 해를 갖는 방정식이 되었다. 문제는 행렬 A 에 선형종속인 열이 있고 $(1, -1)$ 이 영공간 안에 있다는 것이다.

어떤 해 $\hat{x}$ 을 선택해야 할까? 위 그림의 모든 점선은 시간 T 에서 동일한 두 오차 $1, -1$ 을 갖는다. 이러한 오차 $(1, -1) = e = b - p$ 는 가능한 한 작다. 그러나 이 결과가 어떤 점선이 가장 적합한지 알려주지는 않는다.

우선 직감에 따라 높이 2에서 수평선을 선택한다. 만약 최적의 직선에 대한 방정식이 $b = C + Dt$ 라면, 이 경우 선택한 해는 $\hat{x}_1 = C = 2, \hat{x}_2 = D = 0$ 이다. 그러나 만약 직선이 $b = ct + d$ 로 작성되었다면 어떨까? 이 경우도 마찬가지이다(단지 C, D 를 바꾸기만 하면 된다). 이제 수평선은 $\hat{x}_1 = c = 0, \hat{x}_2 = d = 2$ 를 만족한다. 여기서는 도무지 해결책이 보이지 않는다.

4.5절에서는 행렬 A 의 '유사역행렬'을 통해 $A\hat{x} = p$ 의 최단거리 해를 찾는다. 여기서 최단거리 해는 $x^+ = (1, 1)$ 이다. 이는 행렬 A 의 행공간의 특수해이며 길이가 $\sqrt{2}$ 이다(해 $\hat{x} = (2, 0)$ 과 $(0, 2)$ 의 길이는 모두 2이다). 해 x^+ 의 영공간 성분은 0으로 선택한다.

행렬 A 가 선형독립인 열을 갖는 경우, x^+ 에 대한 유사역행렬은 일반적으로 알려진 좌역행렬 $(A^T A)^{-1} A^T$ 이다. 이렇게 표현하면 유사역행렬이 $\hat{x}$ 을 선택하는 가장 좋은 방법이라는 느낌을 준다.

덧붙여, 특이행렬을 이용한 MATLAB 실험에서는 결과가 Inf 또는 NaN(숫자가 아님) 또는 10^{16} (잘못된 숫자)으로 나온다. 모든 경우에 경고가 발생한다. Inf, NaN, 10^{16} 은 각각 $0x = b$, $0x = 0$, $10^{-16}x = 1$ 에서 비롯된다고 본다.

위 3가지는 ‘해가 존재하지 않는 특이행렬, 해가 무수히 많은 특이행렬, 특이행렬에 가까운 행렬’이라는 3가지 난제에 대한 작은 예이다.

포물선 맞추기

공을 던질 때 나타나는 경로를 직선으로 맞추는 것은 어려운 일이다. 포물선 $b = C + Dt + Et^2$ 은 공이 올라갔다 다시 내려오는 경로를 보여준다(여기서 b 는 시간 t 에서의 높이이다). 실제 경로가 완벽한 포물선은 아니지만, 발사체에 대한 전체 이론은 포물선의 근사치에서 출발한다.

갈릴레오(Galileo)가 피사의 사탑에서 돌을 떨어뜨렸을 때, 돌은 가속해서 떨어졌다. 이때 돌의 이동거리에 대한 식에는 2차항 $\frac{1}{2}gt^2$ 이 포함된다(갈릴레오 실험의 핵심은 돌의 질량이 영향을 미치지 않는다는 것이었다). 만약 이 t^2 항이 없었다면 우리는 위성을 궤도로 보낼 수 없었을 것이다. 그러나 t^2 과 같은 비선형함수가 포함되어 있어도, 미지수 C, D, E 는 여전히 선형적으로 나타난다. 점들을 최적의 포물선에 맞추는 것은 여전히 선형대수학의 문제이다. 다음 문제를 생각해 보자.

Q. 포물선 $C + Dt + Et^2$ 에 시간 $t_1, \dots, t_m$ 에서의 높이 $b_1, \dots, b_m$ 을 맞춰라.

A. $m > 3$ 개의 점이 있을 때, 다음과 같이 정확하게 맞춘 m 개의 방정식은 일반적으로 풀 수 없다.

$$\begin{array}{l} C + Dt_1 + Et_1^2 = b_1 \\ \vdots \\ C + Dt_m + Et_m^2 = b_m \end{array} \Leftrightarrow Ax = b \text{ (단, } A = \begin{bmatrix} 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_m & t_m^2 \end{bmatrix} \text{은 } m \times 3 \text{ 행렬)} \quad (10)$$

최소제곱 문제에 대해 생각해 보자. 가장 가까운 포물선 $C + Dt + Et^2$ 을 찾기 위해, 세 정규 방정식 $A^T A \hat{x} = A^T b$ 를 만족하는 $\hat{x} = (C, D, E)$ 를 선택한다.

이 문제를 사영에 대한 문제로 바꿀 수 있는가? 행렬 A 의 열공간의 차원은 ____이다. 벡터 b 의 사영은 $p = A\hat{x}$ 인데, 이는 C, D, E 를 계수로 사용하여 세 열을 결합한 것이다. 첫 번째 데이터 점의 오차는 $e_1 = b_1 - C - Dt_1 - Et_1^2$ 이다. 미적분을 이용해 최소화하는 것을 선호하는 경우, $e_1^2 + \dots + e_m^2$ 의 ____, ____, ____에 대한 편도함수를 구한다. 이 세 도함수를 0으로 둔 방정식이 3×3 연립방정식 $A^T A \hat{x} = A^T b$ 이고, 그 해는 $\hat{x} = (C, D, E)$ 이다.

푸리에 급수(Fourier series)는 무한 차원에서의 최소제곱 문제이다. 이 문제에서는 벡터가 아니라 함수를 근사화한다. 최소화할 함수는 제곱오차의 합에서 제곱오차의 적분으로 변경된다.

EX 3 포물선 $b = C + Dt + Et^2$ 이 시간 $t = 0, 1, 2$ 에서 각각 높이 $b = 6, 0, 0$ 를 지날 때, C, D, E 에 관한 방정식은 다음과 같다.

$$\begin{aligned} C + D \cdot 0 + E \cdot 0^2 &= 6 \\ C + D \cdot 1 + E \cdot 1^2 &= 0 \\ C + D \cdot 2 + E \cdot 2^2 &= 0 \end{aligned} \quad (11)$$

식 (11)이 $Ax = b$ 이고, 이 방정식은 정확하게 해결할 수 있다. 3개의 데이터 점은 3개의 방정식과 정사각행렬을 형성한다. 이때 해는 $x = (C, D, E) = (6, -9, 3)$ 이다. 세 점을 지나는 포물선은 $b = 6 - 9t + 3t^2$ 이다.

4.3 확인문제

4.3A 시간 $t = 1, \dots, 9$ 에서의 9개 측정값 $b_1, \dots, b_9$ 은 모두 0이다. 10번째 측정값 $b_{10} = 40$ 은 이상한(오차가 크게 있는) 측정값이다. 오차 E 에 관한 다음 3가지 옵션에 대해 10개의 점 $(1, 0), (2, 0), \dots, (9, 0), (10, 40)$ 을 맞추기 위한 **최적의 수평선** $y = C$ 를 구하라.

- (a) 오차 제곱합 $E_2 = e_1^2 + \dots + e_{10}^2$ 을 최소화 (그러면 C 에 대한 정규방정식은 선형이다.)
- (b) 최대오차 $E_\infty = |e_{\max}|$ 를 최소화
- (c) 오차 절댓값의 합 $E_1 = |e_1| + \dots + |e_{10}|$ 을 최소화

풀이

- (a) 수평선으로 $0, 0, \dots, 0, 40$ 을 맞추기 위해 오차 제곱합을 최소화하려면 다음과 같이 $C = 40$ 여야 한다.

$$A = (\text{1로 이루어진 열}), \quad A^T A = 10, \quad A^T b = (b_i \text{들의 합}) = 40, \quad \text{따라서 } 10C = 40$$

- (b) 최대오차를 최소화하려면 C 는 0과 40의 중간인 $C = 20$ 이어야 한다.
- (c) 오차 절댓값의 합을 최소화하려면 $C = 0$ 이어야 한다. C 가 0에서 증가할수록 오차 절댓값의 합 $|C| + |40 - C|$ 는 증가한다.

최소 절댓값 합은 중간값(median)에서 나온다(0, 0, ..., 0, 40의 중간값은 0이다). 많은 통계 학자들은 최소제곱법이 $b_{10} = 40$ 과 같은 이상한 측정값에 지나치게 영향을 받는다고 생각하며, 최소 절댓값 합을 선호한다. 그러나 이 경우 방정식은 비선형이다.

이제 (1, 0)부터 (10, 40)까지의 10개 점을 지나는 최소제곱 직선 $C + Dt$ 를 구해보자.

$$A^T A = \begin{bmatrix} 10 & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix} = \begin{bmatrix} 10 & 55 \\ 55 & 385 \end{bmatrix}, \quad A^T b = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix} = \begin{bmatrix} 40 \\ 400 \end{bmatrix}$$

이 값들은 식 (8)에서 나온다. 그러면 $A^T A \hat{x} = A^T b$ 에 따라 $C = -8$, $D = \frac{24}{11}$ 이다.

$b = (0, 0, \dots, 40)$ 에 3을 곱하고 30을 더해 $b_{\text{new}} = (30, 30, \dots, 150)$ 을 얻으면 C 와 D 는 어떻게 되는가? 선형성 때문에 b 를 재조정할 수도 있다. b 에 3을 곱하면 C 와 D 도 3배가 된다. 또한 모든 b_i 에 30을 더하면 C 에도 30이 더해진다.

4.3B 시간 $t = -2, -1, 0, 1, 2$ 에서 값 $b = (0, 0, 1, 0, 0)$ 과 가장 가까운 (최소제곱오차를 갖는) 포물선 $C + Dt + Et^2$ 을 구하라. 먼저 5개의 점을 지나는 포물선의 5개 방정식 $Ax = b$ 를 3개의 미지수 $x = (C, D, E)$ 로 작성하라. 이러한 포물선이 존재하지 않으므로 해는 존재하지 않는다. 그다음 방정식 $A^T A \hat{x} = A^T b$ 를 풀어라.

우선 $D = 0$ 이라고 예측해보자. 최적의 포물선은 왜 $t = 0$ 을 중심으로 대칭이어야 할까? $A^T A \hat{x} = A^T b$ 에서 D 에 대한 두 번째 방정식은 첫 번째, 세 번째 방정식과 분리되어야 한다.

풀이

5개 방정식 $Ax = b$ 는 다음과 같이 직사각행렬인 방데르몽드 행렬(Vandermonde matrix) A 를 갖는다.

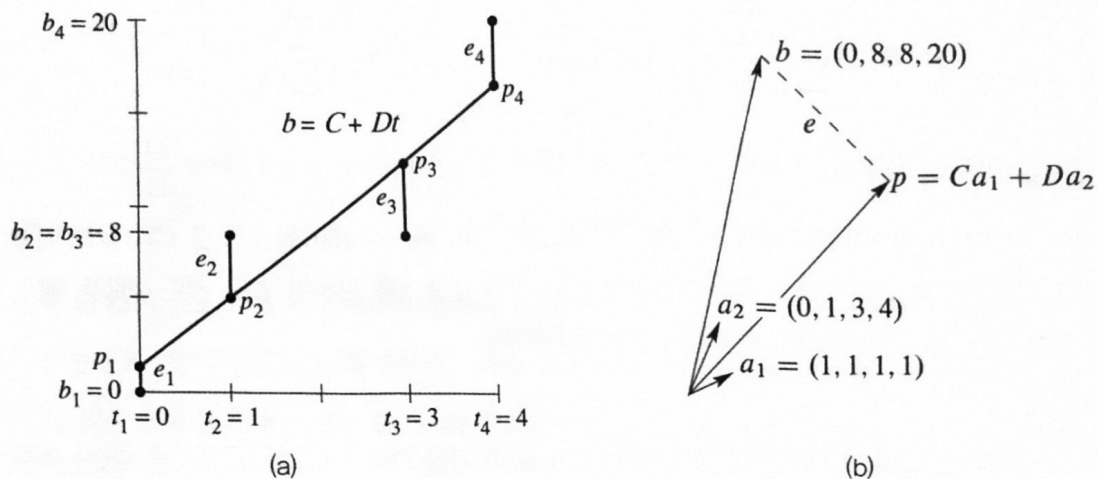
$$\begin{aligned} C + D(-2) + E(-2)^2 &= 0 \\ C + D(-1) + E(-1)^2 &= 0 \\ C + D(0) + E(0)^2 &= 1 \\ C + D(1) + E(1)^2 &= 0 \\ C + D(2) + E(2)^2 &= 0 \end{aligned} \quad A = \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \quad A^T A = \begin{bmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{bmatrix}$$

$A^T A$ 안에 있는 0은 행렬 A 의 2열이 1열, 3열과 직교하기 때문에 생긴 것이다. 이는 행렬 A 를 통해 바로 확인할 수 있다(시간 $-2, -1, 0, 1, 2$ 는 서로 대칭). 다음과 같이 포물선 $C + Dt + Et^2$ 에서의 최적의 C, D, E 를 $A^T A \hat{x} = A^T b$ 로부터 찾을 수 있고, D 는 C, E 와 관련이 없다.

$$\begin{bmatrix} 5 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{aligned} C &= \frac{34}{70} \\ D &= 0 \text{ (예상대로)} \\ E &= -\frac{10}{70} \end{aligned}$$

4.3 연습문제

※ [문제 1~11]은 4개의 데이터 점 $b = (0, 8, 8, 20)$ 을 이용한 문제이다. 다음 그림을 참고한다.



[그림 4.8] [문제 1~11] 가장 가까운 직선 $C + Dt$ 는 $\mathbb{R}^4$ 에서 $Ca_1 + Da_2$ 에 해당한다.

1. 시간 $t = 0, 1, 3, 4$ 에서 $b = 0, 8, 8, 20$ 일 때, 정규방정식 $A^T A \hat{x} = A^T b$ 를 만들고 풀어라. [그림 4.8(a)]에서 최적의 직선에 대해 4개의 높이 p_i 와 4개의 오차 e_i 를 구하라. $E = e_1^2 + e_2^2 + e_3^2 + e_4^2$ 의 최솟값은 무엇인가?
2. (직선 $C + Dt$ 는 p_i 들을 지난다.) 시간 $t = 0, 1, 3, 4$ 에서 측정값이 $b = 0, 8, 8, 20$ 일 때, 4개의 방정식 $Ax = b$ 를 작성하라(단, 이 방정식의 해는 존재하지 않는다). 측정값을 $p = 1, 5, 13, 17$ 로 바꾸고 $A\hat{x} = p$ 의 정확한 해를 구하라.
3. $e = b - p = (-1, 3, -5, 3)$ 이 행렬 A 의 두 열과 직교하는지 확인하라. 벡터 b 로부터 행렬 A 의 열공간까지의 최단거리 $\|e\|$ 를 구하라.
4. (미분적분학을 이용해) $E = \|Ax - b\|^2$ 을 4개 제곱항의 합으로 표현하라. 마지막 항은 $(C + 4D - 20)^2$ 이다. 편미분방정식 $\frac{\partial E}{\partial C} = 0, \frac{\partial E}{\partial D} = 0$ 을 구하라. 그리고 이 식의 양변을 2로 나누어 정규방정식 $A^T A \hat{x} = A^T b$ 를 구하라.
5. $b = (0, 8, 8, 20)$ 과 가장 가까운 최적의 수평선의 높이 C 를 구하라. 점을 수평선에 정확하게 맞추려면 풀 수 없는 방정식 $C = 0, C = 8, C = 8, C = 20$ 을 풀어야 한다. 4×1 행렬 A 를 구하고 방정식 $A^T A \hat{x} = A^T b$ 를 풀어라. 높이 $\hat{x} = C$ 에서의 수평선과 e 에서의 4개의 오차를 그려라.

6. $b = (0, 8, 8, 20)$ 을 $a = (1, 1, 1, 1)$ 방향 직선 위로 사영하라. $\hat{x} = \frac{a^T b}{a^T a}$ 와 사영 $p = \hat{x}a$ 를 구하라. $e = b - p$ 가 a 와 직교하는지 확인하고, b 에서부터 a 방향 직선까지의 최단거리 $\|e\|$ 를 구하라.
7. 4개의 점 $b = (0, 8, 8, 20)$ 을 가장 가깝게 지나는 직선 $b = Dt$ 를 찾으라. 정확하게 지나는 직선은 $D \cdot 0 = 0, D \cdot 1 = 8, D \cdot 3 = 8, D \cdot 4 = 20$ 을 만족해야 한다. 4×1 행렬 A 를 구하고 방정식 $A^T A \hat{x} = A^T b$ 를 풀어라. 최적의 직선 $b = Dt$ 와 오차 e_i 들을 나타내서 [그림 4.8(a)]를 다시 그려라.
8. $b = (0, 8, 8, 20)$ 을 $a = (0, 1, 3, 4)$ 방향 직선 위로 사영하라. $\hat{x} = D$ 와 $p = \hat{x}a$ 를 구하라. [문제 5]~[문제 6]에서의 최적의 C 와 [문제 7]~[문제 8]에서의 최적의 D 는 [문제 1]~[문제 4]에서의 최적의 (C, D) 와 일치하지 않는다. 그 이유는 $(1, 1, 1, 1)$ 과 $(0, 1, 3, 4)$ 가 _____한 관계이기 때문이다.
9. 4개의 점 $b = (0, 8, 8, 20)$ 을 가장 가깝게 지나는 포물선 $b = C + Dt + Et^2$ 에 대해 3개의 미지수 $x = (C, D, E)$ 를 갖는 방정식 $Ax = b$ 를 작성하라(단, 이 방정식의 해는 존재하지 않는다). 3개의 정규방정식 $A^T A \hat{x} = A^T b$ 를 작성하라(단, 해는 구하지 않는다). [그림 4.8(a)]에서는 4개의 점을 포물선에 맞추고 있다. 그렇다면 [그림 4.8(b)]는 어떻게 되는가?
10. 4개의 점 $b = (0, 8, 8, 20)$ 을 가장 가깝게 지나는 3차곡선 $b = C + Dt + Et^2 + Ft^3$ 에 대해 4개의 방정식 $Ax = b$ 를 작성하라. 그리고 소거법을 이용해 이 방정식을 풀어라. [그림 4.8(a)]에서 3차함수는 4개의 점을 모두 지난다. p 와 e 를 모두 구하라.
11. 4개 시간의 평균은 $\hat{t} = \frac{1}{4}(0+1+3+4) = 2$ 이다. 또한 4개 점 b 의 평균은 $\hat{b} = \frac{1}{4}(0+8+8+20) = 9$ 이다. 다음 물음에 답하라.
- (a) 최적의 직선이 평균점 $(\hat{t}, \hat{b}) = (2, 9)$ 를 지남을 확인하라.
- (b) 왜 $C + D\hat{t} = \hat{b}$ 이 $A^T A \hat{x} = A^T b$ 의 첫 번째 방정식으로부터 나오는지 설명하라.

※ [문제 12~16]은 통계학의 기본 아이디어(최소제곱법의 기초)를 소개한다.

12. (꼭 풀어볼 만한 문제) 이 문제에서는 $b = (b_1, \dots, b_m)$ 을 $a = (1, \dots, 1)$ 방향 직선 위로 사영한다. (최소제곱법을 이용해) 1개의 미지수가 있는 m 개의 방정식 $ax = b$ 를 풀고자 한다. 다음 물음에 답하라.
- (a) $a^T a \hat{x} = a^T b$ 를 풀어서 $\hat{x}$ 이 b 의 성분들의 평균임을 보여라.
- (b) 오차 $e = b - a\hat{x}$ 와 분산 $\|e\|^2$, 표준편차 $\|e\|$ 를 구하라.
- (c) 수평선 $\hat{b} = 3$ 은 $b = (1, 2, 6)$ 과 가장 가깝다. $p = (3, 3, 3)$ 이 e 와 직교하는지 확인하고, 3×3 사영행렬 P 를 구하라.

13. 최소제곱법에 대한 첫 번째 가정은 $Ax = b$ (평균이 0인 잡음 e)이다. 오차벡터 $e = b - Ax$ 의 양변에 $(A^T A)^{-1} A^T$ 를 곱하여 우변에 $\hat{x} - x$ 를 만들어라. 추정오차 $\hat{x} - x$ 의 평균은 0이다. 따라서 추정 $\hat{x}$ 은 편향되지 않는다.
14. 최소제곱법에 대한 두 번째 가정은 m 개의 오차 e_i 가 분산 σ^2 을 가지며 선형독립이라는 것이다. 따라서 $(b - Ax)(b - Ax)^T$ 의 평균은 $\sigma^2 I$ 이다. 이 식의 양변 왼쪽에는 $(A^T A)^{-1} A^T$ 를, 오른쪽에는 $A(A^T A)^{-1}$ 를 곱하여 평균행렬 $(\hat{x} - x)(\hat{x} - x)^T$ 가 $\sigma^2 (A^T A)^{-1}$ 임을 보여라. 이는 10.3절에서 살펴볼 공분산행렬(covariance matrix) W 이다.
15. 의사가 심박수를 4회 측정했다고 하자. $x = b_1, \dots, x = b_4$ 에 대한 최적의 해는 $b_1, \dots, b_4$ 의 평균 $\hat{x}$ 이다. 행렬 A 는 열이 1개이고 모든 성분이 1이다. [문제 14]에서는 추정오차 $(\hat{x} - x)^2$ 이 $\sigma^2 (A^T A)^{-1} = \underline{\hspace{2cm}}$ (으)로 주어졌다. 평균을 취하면 분산은 σ^2 에서 $\frac{\sigma^2}{4}$ 으로 감소한다.
16. 9개의 수 $b_1, \dots, b_9$ 의 평균 $\hat{x}_9$ 을 안다고 할 때, 하나의 수 b_{10} 을 추가하면 10개 수의 평균 $\hat{x}_{10}$ 을 어떻게 빠르게 구할 수 있는가? 재귀적 최소제곱법의 아이디어는 10개 수를 모두 더하지 않는 것이다. 다음과 같이 $\hat{x}_{10}$ 을 구할 때, $\hat{x}_9$ 에 곱해진 값은 무엇인가?

$$\hat{x}_{10} = \frac{1}{10} b_{10} + \underline{\hspace{2cm}} \hat{x}_9 = \frac{1}{10} (b_1 + \dots + b_{10}) \text{ ([확인문제 4.2C] 참고)}$$

※ [문제 17~24]를 통해 $\hat{x}$, p , e 에 대해 더 많이 연습할 수 있다.

17. 직선 $b = C + Dt$ 가 $t = -1$ 에서 $b = 7$ 을 지나고, $t = 1$ 에서 $b = 7$ 을 지나며, $t = 2$ 에서 $b = 21$ 을 지나도록 하는 세 방정식을 작성하라. 그리고 최소제곱해 $\hat{x} = (C, D)$ 를 구한 뒤 가장 가까운 직선을 그려라.
18. [문제 17]에서 사영 $p = A\hat{x}$ 을 구하라. 이를 통해 최적의 직선의 높이 3개를 알 수 있다. 오차벡터가 $e = (2, -6, 4)$ 임을 보여라. 왜 $Pe = 0$ 인가?
19. 시간 $t = -1, 1, 2$ 에서의 측정값이 [문제 18]에서의 오차 2, -6, 4와 같다고 가정하자. $\hat{x}$ 을 계산하고, 이 측정값에 대한 최적의 직선을 구하라. 그리고 작성한 답에 대해 다음 설명의 빈칸을 작성하라.
- $b = (2, -6, 4)$ 는 $\underline{\hspace{2cm}}$ 와(과) 직교한다. 따라서 사영은 $p = 0$ 이다.

20. 시간 $t = -1, 1, 2$ 에서의 측정값이 $b = (5, 13, 17)$ 이다. $\hat{x}$ 을 계산하고, 최적의 직선과 e 를 구하라. 오차는 $e = 0$ 이다. 왜냐하면 b 는 _____이기 때문이다.
21. 4개의 기본 부분공간 중 무엇이 오차벡터 e 를 포함하는가? 또 무엇이 p 를 포함하며, 무엇이 $\hat{x}$ 을 포함하는가? 그리고 행렬 A 의 영공간은 무엇인가?
22. 시간 $t = -2, -1, 0, 1, 2$ 일 때, $b = 4, 2, -1, 0, 0$ 과 가장 가까운 최적의 직선 $C + Dt$ 를 구하라.
23. 오차벡터 e 는 b 또는 p 또는 e 또는 $\hat{x}$ 과 직교하는가? $\|e\|^2$ 이 $e^T b$, 즉 $b^T b - p^T b$ 와 동일함을 보여라. 이는 최소의 전체 오차 E 이다.
24. $x_1, \dots, x_n$ 에 대한 $\|Ax\|^2$ 의 편도함수는 벡터 $2A^T Ax$ 가 된다. $2b^T Ax$ 의 도함수들은 벡터 $2A^T b$ 가 된다. 따라서 $\|Ax - b\|^2$ 의 도함수는 _____일 때 0이 된다.

도전문제

25. 어떤 조건을 만족할 때, 세 점 $(t_1, b_1), (t_2, b_2), (t_3, b_3)$ 가 한 직선 위에 놓이는가? 열공간에 대한 답을 말하자면, (b_1, b_2, b_3) 는 $(1, 1, 1)$ 과 (t_1, t_2, t_3) 의 선형결합이어야 한다. t_i 들과 b_i 들을 연결하는 특별한 방정식을 만들어라.
26. 4개의 모서리 $(1, 0), (0, 1), (-1, 0), (0, -1)$ 에서 4개의 값 $b = (0, 1, 3, 4)$ 를 갖는 평면을 찾으라. 4개의 점에서 방정식 $C + Dx + Ey = b$ 는 3개의 미지수 $x = (C, D, E)$ 가 있는 $Ax = b$ 이다. 행렬 A 는 무엇인가? 정사각형의 중심 $(0, 0)$ 에서 $C + Dx + Ey = (b_i \text{들의 평균})$ 이 됨을 보여라.
27. (직선 간의 거리 문제) 두 점 $P = (x, x, x)$ 와 $Q = (y, 3y, -1)$ 은 공간상에서 만나지 않는 두 직선 위에 있다. x 와 y 를 잘 선택해서 제곱거리 $\|P - Q\|^2$ 을 최소화하라. 가장 가까운 P 와 Q 를 연결하는 직선은 _____와(과) 직교한다.
28. 행렬 A 의 열들이 선형독립이 아니라고 가정하자. $P = B(B^T B)^{-1} B^T$ 가 행렬 A 의 열공간 위로 사영되도록 하는 행렬 B 를 구하라($A^T A$ 가 가역행렬이 아니면 일반적인 공식 $P = A(A^T A)^{-1} A^T$ 는 실패한다).

29. 보통 $\mathbb{R}^n$ 에서 n 개의 점 $x=0, a_1, \dots, a_{n-1}$ 을 포함하는 초평면(hyperplane)은 오직 하나이다(예를 들어 $n=3$ 일 때, _____인 경우 $0, a_1, a_2$ 를 포함하는 평면은 오직 하나이다). $\mathbb{R}^n$ 에서 정확히 하나의 평면을 갖는지를 판정하는 방법은 무엇인가?

30. [EX 2]에서는 시간 t_i 들을 이동시켜 이들의 합이 0이 되게 만들었다. 즉 평균 시간 $\hat{t} = \frac{t_1 + \dots + t_m}{m}$ 을 뺀으로써 $T_i = t_i - \hat{t}$ 을 얻는다. 이제 열 $(1, \dots, 1)$ 과 $(T_1, \dots, T_m)$ 은 직교하며, $A^T A$ 는 대각행렬이다. $A^T A$ 의 성분은 m 과 $T_1^2 + \dots + T_m^2$ 이다. 최적의 C, D 에 대한 공식이 다음과 같음을 보여라.

$$T = t - \hat{t}, \quad C = \frac{b_1 + \dots + b_m}{m}, \quad D = \frac{b_1 T_1 + \dots + b_m T_m}{T_1^2 + \dots + T_m^2}$$

최적의 직선은 $C + DT$ 또는 $C + D(t - \hat{t})$ 이다. $A^T A$ 를 대각행렬로 만드는 시간 이동은 그람-슈미트 과정(Gram-Schmidt process)의 예이다. 이 과정에서는 **행렬 A 의 열들을 미리 직교화한다**. 이와 관련한 내용은 4.4절에서 자세히 살펴본다.

핵심 아이디어 리뷰

- ❶ 최소제곱해 $\hat{x}$ 은 $\|Ax - b\|^2 = x^T A^T A x - 2x^T A^T b + b^T b$ 를 최소화한다. 이는 m 개의 방정식에서의 오차 $e_1, \dots, e_m$ 의 제곱의 합인 E 이다(단, $m > n$).
- ❷ 최적의 $\hat{x}$ 은 정규방정식 $A^T A \hat{x} = A^T b$ 로부터 구한다. E 는 최솟값이다.
- ❸ 직선 $b = C + Dt$ 에 m 개의 점을 최대한 맞추려면 C, D 는 정규방정식을 만족해야 한다.
- ❹ 최적의 직선의 높이는 $p = (p_1, \dots, p_m)$ 이다. 또한 데이터 점까지의 세로방향 거리는 오차 $e = (e_1, \dots, e_m)$ 이다. 핵심 방정식은 $A^T e = 0$ 이다.
- ❺ m 개의 점을 $n (< m)$ 개의 함수의 선형결합으로 표현하려고 하면, 일반적으로 m 개의 방정식 $Ax = b$ 는 풀 수 없다. n 개의 방정식 $A^T A \hat{x} = A^T b$ 는 최소제곱해를 준다. 여기서 최소제곱해는 최소 **평균제곱오차(MSE; Mean Square Error)**의 선형결합이다.

4.4 정규직교기저와 그람-슈미트 과정

1. 열벡터 $q_1, \dots, q_n$ 은 $q_i^T q_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$ 일 때 정규직교(orthonormal)한다고 한다. 그러면 $Q^T Q = I$ 이다.
2. Q 가 정사각행렬이면 $QQ^T = I$ 이고 $Q^T = Q^{-1}$ 이다. 그러면 Q 를 직교행렬(orthogonal matrix)이라 한다.
3. $Qx = b$ 의 최소제곱해는 $\hat{x} = Q^T b$ 이다. b 의 사영은 $p = QQ^T b = Pb$ 이다.
4. 그람-슈미트 과정(Gram-Schmidt process)은 선형독립인 열 a_i 를 정규직교하는 열 q_i 로 변환한다. 먼저 $q_1 = \frac{a_1}{\|a_1\|}$ 으로 시작한다.
5. q_i 는 $\frac{a_i - (a_i \text{의 사영 } p_i)}{\|a_i - p_i\|}$ 이다. 또한 사영 p_i 는 $p_i = (a_i^T q_1)q_1 + \dots + (a_i^T q_{i-1})q_{i-1}$ 이다.
6. 각 a_i 는 $q_1, \dots, q_i$ 의 선형결합이다. 그러면 $A = QR$ 이다(여기서 Q 는 직교행렬, R 은 삼각행렬).

이 절의 목표는 2가지이다. 첫 번째 목표는 행렬 A 에서 직교하는 열들이 좋은 이유를 이해하는 것이다. 직교하는 열들의 내적은 0이므로, $A^T A$ 는 대각행렬이 된다. 따라서 $\hat{x}, p = A\hat{x}$ 을 아주 쉽게 구할 수 있다. 두 번째 목표는 직교벡터 q_i 를 만드는 것이다. 이 절에서는 그람-슈미트 과정이 어떻게 원래 기저벡터 a_i 를 직교벡터 q_i 로 변환하는지 확인할 수 있다. a_i 는 행렬 A 의 열이고, 이들은 일반적으로 서로 직교하지 않는다. 정규직교 기저벡터 q_i 는 새로운 행렬 Q 의 열이 된다.

3장에 따르면, 기저는 벡터공간을 생성하는 선형독립인 벡터들로 구성된다. 기저벡터들은 $0^\circ, 180^\circ$ 만 제외하면 어떤 각으로든 만날 수 있다. 그러나 기저벡터의 방향을 x, y, z 축으로 잡으면 이들은 항상 직교한다. 좌표축은 편리함을 이유로 항상 직교한다고 생각한다. 이는 기저벡터와 관련한 그림과 계산을 모두 간단하게 만든다.

내적 $q_i \cdot q_j$ 가 모두 0일 때 $q_1, \dots, q_n$ 은 서로 직교한다. 좀 더 정확히 말하면 $i \neq j$ 일 때만 $q_i^T q_j = 0$ 이다. 한발 더 나아가 각 벡터 $q_1, \dots, q_n$ 들을 자기 자신의 크기로 나누면 정규직교 벡터가 된다. 이를 통해 모든 벡터의 크기는 1이 된다.

정의 정규직교

n 개의 벡터 $q_1, \dots, q_n$ 이 다음을 만족할 때, 이 벡터들은 **정규직교**한다고 한다.

$$q_i^T q_j = \begin{cases} 0, & i \neq j \text{ (직교벡터)} \\ 1, & i = j \text{ (단위벡터 : } \|q_i\| = 1) \end{cases}$$

정규직교하는 열을 가진 행렬 Q 는 $Q^T Q = I$ 를 만족한다. 대부분의 경우 Q 는 직사각행렬 ($m > n$)이다.

$Q^T Q = I$ 이므로 행렬 Q 는 다루기가 쉽다. 행렬 언어로 표현하면, 열 $q_1, \dots, q_n$ 은 정규직교한다. Q 가 정사각행렬일 필요는 없으므로 일반적으로 $QQ^T \neq I$ 이다.

정규직교하는 열을 가진 행렬 Q 는 다음과 같이 $Q^T Q = I$ 를 만족한다.

$$Q^T Q = \begin{bmatrix} -q_1^T - \\ -q_2^T - \\ \vdots \\ -q_n^T - \end{bmatrix} \begin{bmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I \quad (1)$$

Q^T 의 i 행과 Q 의 j 열의 내적은 $q_i^T q_j$ 이다. $Q^T Q$ 의 주대각성분이 아닐 경우(즉 $i \neq j$), 직교성에 따라 내적은 0이다. 주대각성분인 경우(즉 $i = j$), 각 벡터는 정규벡터이므로 $q_i^T q_i = \|q_i\|^2 = 1$ 이다. 대부분의 경우 Q 는 직사각행렬이다($m > n$). 종종 $m = n$ 이다. Q 가 정사각행렬일 때, $Q^T Q = I$ 이면 $Q^T = Q^{-1}$ 이다. 즉 (전치행렬) = (역행렬)이다.

열들이 정규성은 만족하지 않고 직교성만 만족하는 경우, 내적은 여전히 대각행렬(항등행렬은 아님)을 만든다. 이 대각행렬은 항등행렬 I 만큼 좋은 행렬이다. 여기서 중요한 것은 직교성이다. 정규화, 즉 벡터의 크기를 1로 만드는 것은 쉽다.

반복해서 말하지만 Q 가 직사각행렬이라고 해도 $Q^T Q = I$ 는 성립한다. Q^T 는 왼쪽에 곱해지는 한쪽 방향 역행렬이다. 만약 Q 가 정사각행렬이면 $QQ^T = I$ 도 성립한다. 이때 Q^T 는 Q 의 양쪽 방향 역행렬이다. 정사각행렬 Q 의 행도 열처럼 정규직교한다. Q 의 역행렬은 전치행렬 Q^T 이다. 이와 같은 경우 정사각행렬 Q 를 **직교행렬**이라 한다.³

3 '정규직교행렬(orthonormal matrix)'이 Q 에 더 적합한 이름일 수 있지만, 그렇게 부르지는 않는다. 열들이 정규직교하는 행렬은 언제나 Q 로 표기한다. 그러나 Q 가 정사각행렬인 경우에만 '직교행렬(orthogonal matrix)'이라고 한다.

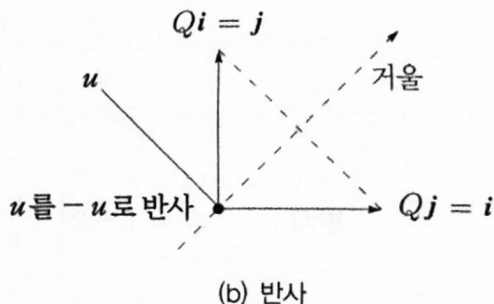
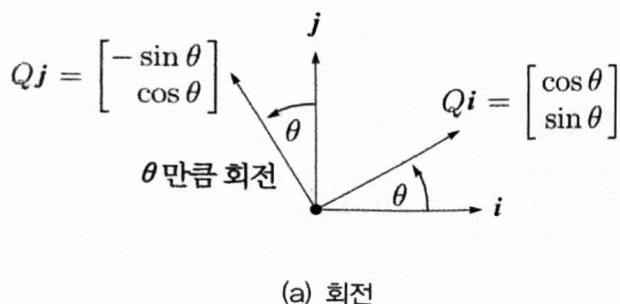
이제 직교행렬의 3가지 예(회전, 치환, 반사)를 살펴보자. 이들은 모두 중요한 행렬이다. 직교행렬인지를 빠르게 판정하는 방법은 $Q^T Q = I$ 인지 확인하는 것이다.

EX 1 (회전) 다음과 같은 행렬 Q 는 평면 위에 있는 모든 벡터를 각도 θ 만큼 회전한다.

$$Q = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, \quad Q^T = Q^{-1} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

행렬 Q 의 열들은 서로 직교한다(내적을 계산해보자). $\sin^2 \theta + \cos^2 \theta = 1$ 이므로 열들의 크기는 모두 1이다. 이 열들은 공간 $\mathbb{R}^2$ 의 정규직교기저를 이룬다.

[그림 4.9(a)]에서 표준기저벡터 i 와 j 가 θ 만큼 회전한다. 그러면 Q^{-1} 는 벡터를 $-\theta$ 만큼 회전한다. $-\theta$ 와 θ 의 코사인 값이 동일하고 $\sin(-\theta) = -\sin \theta$ 이므로, Q^{-1} 는 Q^T 와 같다. 따라서 $Q^T Q = I$ 이고 $Q Q^T = I$ 이다.



[그림 4.9] $Q = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ 에 의한 회전과 $Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 에 의한 45° 직선 기준의 반사

EX 2 (치환) 다음 행렬들은 (x, y, z) 의 순서를 (y, z, x) 로 바꾸고, (x, y) 의 순서를 (y, x) 로 바꾼다.

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ z \\ x \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$$

행렬 Q 의 모든 열은 크기가 1이고, 서로 직교한다(성분 1이 서로 다른 위치에 있기 때문). 따라서 치환행렬 Q 의 역행렬은 전치행렬과 같다($Q^{-1} = Q^T$). 역행렬은 다음과 같이 벡터의 성분들을 원래 위치로 되돌려 놓는다.

$$(\text{역행렬}) = (\text{전치행렬}) \quad \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y \\ z \\ x \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

치환행렬에 대해 다음이 성립한다.

모든 치환행렬은 직교행렬이다.

EX 3 (반사) u 가 임의의 단위벡터(크기가 1인 벡터)이면 $Q = I - 2uu^T$ 로 놓는다. $u^T u$ 는 수 $\|u\|^2 = 1$ 이 되지만, uu^T 는 랭크 1 행렬이다. 그러면 Q^T 와 Q^{-1} 는 다음과 같이 모두 Q 와 동일하다.

$$Q^T = I - 2uu^T = Q, \quad Q^T Q = I - 4uu^T + 4uu^T uu^T = I \quad (2)$$

반사행렬 $Q = I - 2uu^T$ 는 대칭행렬인 동시에 직교행렬이다. 이 행렬을 제곱하면 항등행렬이 된다($Q^2 = Q^T Q = I$). $(-1)^2 = 1$ 이 되는 것처럼, 거울에 반사를 두 번 하면 원래대로 돌아온다. 식 (2)의 $4uu^T uu^T$ 에서 $u^T u = 1$ 이 됨을 주목하라.

반사의 기준의 예를 $u = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ 로 들어 살펴보자. $2uu^T$ (열과 행의 곱)를 계산하고, I 에서 $2uu^T$ 를 뺌으로써 다음과 같이 반사행렬 Q 를 얻는다.

$$\text{반사} \quad Q = I - 2 \begin{bmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$$

(x, y) 는 반사행렬에 의해 (y, x) 로 움직이지만, $(3, 3)$ 은 움직이지 않는다. 왜냐하면 $(3, 3)$ 은 반사 직선 위에 있기 때문이다.

회전은 모든 벡터의 크기를 유지한다. 반사와 치환도 마찬가지이다. 그러므로 행렬 Q 를 곱한다고 해도 벡터의 크기와 벡터 간의 사잇각은 변하지 않는다. 이를 간략히 증명하면 $(Qx)^T(Qx) = x^T Q^T Q x = x^T I x = x^T x$ 이므로, $\|Qx\|^2$ 은 $\|x\|^2$ 과 같음을 알 수 있다.

행렬 Q 가 정규직교하는 열을 가지면($Q^T Q = I$), Q 를 곱하더라도 다음과 같이 크기가 변하지 않는다.

$$\text{모든 벡터 } x \text{에 대해 } \|Qx\| = \|x\| \quad (3)$$

Q 는 다음과 같이 내적 또한 보존한다. 다음 식에서는 단지 $Q^T Q = I$ 만을 이용했다.

$$(Qx)^T(Qy) = x^T Q^T Q y = x^T y$$

정규직교기저를 이용한 사영 QQ^T : Q 가 A 를 대체하는 경우

직교행렬은 계산에 유리하다. 왜냐하면 벡터의 크기가 고정되므로 수가 너무 커지는 경우가 없기 때문이다. 안정적인 컴퓨터 코드는 Q 를 가능한 한 많이 사용한다.

부분공간 위로의 사영에서 모든 공식은 $A^T A$ 를 포함한다. $A^T A$ 의 성분들은 기저벡터 $a_1, \dots, a_n$ 의 내적 $a_i^T a_j$ 이다.

기저벡터들이 정규직교한다고 가정하면 벡터 a_i 들은 q_i 들로 나타낼 수 있다. 그러면 $A^T A$ 는 $Q^T Q = I$ 로 간단해진다. $\hat{x}$, p , P 는 다음과 같이 바뀐다.

$$\hat{x} = Q^T b, \quad p = Q\hat{x}, \quad P = QQ^T \quad (4)$$

$Qx = b$ 의 최소제곱해는 $\hat{x} = Q^T b$ 이다. 그리고 사영행렬은 QQ^T 이다.

여기서는 역행렬을 구할 필요가 없다. 이것이 바로 정규직교기저를 사용하는 이유이다. 최적의 해 $\hat{x} = Q^T b$ 는 $q_1, \dots, q_n$ 과 b 의 내적이다. 이를 통해 1차원 사영을 얻는다. **결합행렬(coupling matrix)** 또는 **상관행렬(correlation matrix)**인 $A^T A$ 는 이제 $Q^T Q = I$ 가 된다. 이로써 결합(coupling)은 사라진다. 즉 A 가 정규직교하는 열을 가진 행렬 Q 이면, 사영은 $p = Q\hat{x} = QQ^T b$ 이다.

$$q_i \text{ 위로의 사영} \quad p = \begin{bmatrix} | & & | \\ q_1 & \cdots & q_n \\ | & & | \end{bmatrix} \begin{bmatrix} q_1^T b \\ \vdots \\ q_n^T b \end{bmatrix} = q_1(q_1^T b) + \cdots + q_n(q_n^T b) \quad (5)$$

중요한 경우를 하나 생각해보자. Q 가 정사각행렬($m = n$ 인 경우)이면, 부분공간은 전체 공간이 된다. 그러면 $Q^T = Q^{-1}$ 이고 $\hat{x} = Q^T b$ 는 $x = Q^{-1}b$ 와 같다. 이때 해는 근사해가 아닌 정확한 해이다. 벡터 b 의 전체 공간 위로의 사영은 b 자기 자신이다. 이 경우 $p = b$, $P = QQ^T = I$ 가 된다.

전체 공간 위로의 사영은 굳이 언급할 가치가 없다고 생각할 수도 있다. 그러나 $p = b$ 일 때, 우리의 공식은 b 를 1차원 사영들을 모아 놓은 것으로 조합한다. $q_1, \dots, q_n$ 이 전체 공간에 대한 정규직교 기저벡터인 경우, Q 는 정사각행렬이다. 모든 $b = QQ^T b$ 는 q_i 들을 따라 사영한 것들의 합이다.

$$b = q_1(q_1^T b) + q_2(q_2^T b) + \cdots + q_n(q_n^T b) \quad (6)$$

이제 변환에 대해 알아보자. $QQ^T = I$ 는 푸리에 급수와 응용수학의 모든 위대한 변환에 대한 기초가 된다. 변환은 벡터 b 또는 함수 $f(x)$ 를 직교하는 조각들로 분해한다. 그다음 식 (6)에서와

같이 조각들을 합침으로써, 역변환은 b 또는 $f(x)$ 를 원래대로 되돌린다.

EX 4 다음과 같이 행렬 Q 의 직교하는 열들이 정규직교벡터 q_1, q_2, q_3 이다.

$$m = n = 3, \quad Q = \frac{1}{3} \begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}, \quad Q^T Q = Q Q^T = I$$

$b = (0, 0, 1)$ 을 각각 따로 q_1, q_2, q_3 위로 사영하면 p_1, p_2, p_3 이다.

$$q_1(q_1^T b) = \frac{2}{3}q_1, \quad q_2(q_2^T b) = \frac{2}{3}q_2, \quad q_3(q_3^T b) = -\frac{1}{3}q_3$$

처음 2개의 합은 벡터 b 를 q_1, q_2 가 포함된 평면 위로 사영한 것이다. 3개 모두의 합은 벡터 b 를 전체 공간 위로 사영한 것으로, 이는 $p_1 + p_2 + p_3 = b$ 자기 자신이다.

$$\begin{array}{l} \text{b의 재구성} \\ b = p_1 + p_2 + p_3 \end{array} \quad \frac{2}{3}q_1 + \frac{2}{3}q_2 - \frac{1}{3}q_3 = \frac{1}{9} \begin{bmatrix} -2+4-2 \\ 4-2-2 \\ 4+4+1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = b$$

그람-슈미트 과정

이 절의 핵심은 ‘직교성이 좋다’는 사실을 이해하는 것이다. 사영과 최소제곱법은 $A^T A$ 와 관련이 있다. $A^T A$ 가 $Q^T Q = I$ 가 되면 역행렬을 계산할 필요가 없다. 1차원 사영들은 서로 간의 간섭이 없다. 최적의 $\hat{x}$ 은 $Q^T b$ 이다(이는 단지 n 개의 분리된 내적이다). 이렇게 되려면 ‘벡터들은 서로 정규직교’해야 한다. 이제 정규직교벡터들을 만들기 위한 ‘그람-슈미트 과정’에 대해 알아보자.

선형독립인 벡터 a, b, c 로 시작하자. 직교하는 세 벡터 A, B, C 를 만들고자 한다. 마지막으로 A, B, C 를 각각의 크기로 나눈다. 이를 통해 세 정규직교벡터 $q_1 = \frac{A}{\|A\|}, q_2 = \frac{B}{\|B\|}, q_3 = \frac{C}{\|C\|}$ 를 만들 수 있다.

그람-슈미트 과정은 $A = a$ 로 놓고 시작한다. 첫 번째 방향은 그대로 받아들인다. 다음 방향 B 는 A 와 직교해야 한다. 이를 위해 다음과 같이 b 에서 b 의 A 위로의 사영을 뺀다. 그러면 A 와 직교하는 부분이 남는데, 이것이 바로 벡터 B 이다.

그람-슈미트의 첫 번째 단계
 B 는 A 와 직교

$$B = b - \frac{A^T b}{A^T A} A \quad (7)$$

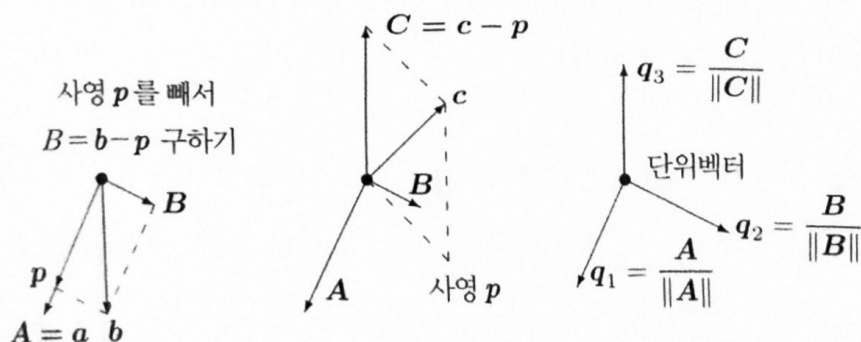
[그림 4.10]에서 A 와 B 는 직교한다. 식 (7)의 양변에 A^T 를 곱하여 $A^T B = A^T b - A^T b = 0$ 이 됨을 확인하라. 벡터 B 는 소위 말하는 오차벡터 e 로, e 는 A 와 직교한다. 식 (7)에서 B 는 0이 아니다(만약 0이 아니면, a 와 b 는 선형종속이므로 모순이다). 이제 A 와 B 의 방향이 정해졌다.

세 번째 방향은 c 로부터 온다. 세 번째 방향은 A 와 B 의 선형결합이 아니다. 왜냐하면 c 는 a 와 b 의 선형결합이 아니기 때문이다. 대부분의 경우 c 는 A , B 와 직교하지 않는다. 그러므로 다음과 같이 c 에서 A , B 방향의 성분을 뺀으로써 C 를 얻는다.

그람-슈미트의 다음 단계
 C 는 A , B 와 직교

$$C = c - \frac{A^T c}{A^T A} A - \frac{B^T c}{B^T B} B \quad (8)$$

지금까지 살펴본 내용이 바로 그람-슈미트 과정의 아이디어이다. 새로운 벡터를 기존에 구한 직교하는 벡터들 위로 사영하여, 그 성분들을 새로운 벡터에서 뺀다. 이 과정을 매번 반복한다.⁴ 만약 네 번째 벡터 d 가 있다면, d 에서 A , B , C 위로의 3개의 사영을 뺀으로써 D 를 구한다. 마지막으로 직교벡터 A , B , C , D 를 각각의 크기로 나눈다. 그 결과로 얻는 q_1 , q_2 , q_3 , q_4 는 정규직교벡터이다.



[그림 4.10] 먼저 a 를 지나는 직선 위로 b 를 사영하여 $B = b - p$ 를 구한다. 그다음 c 를 AB -평면 위의 p 로 사영하여 $C = c - p$ 를 구한다. 구한 A , B , C 를 각각 $\|A\|$, $\|B\|$, $\|C\|$ 로 나눈다.

⁴ 그람(Gram)이 이 아이디어를 생각했던 것 같다. 여기서 슈미트(Schmidt)의 역할은 알 수 없다.

그람-슈미트 과정에 관한 예를 하나 살펴보자. 다음 벡터 a, b, c 는 선형독립이고 직교하지 않는다.

$$a = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}, \quad c = \begin{bmatrix} 3 \\ -3 \\ 3 \end{bmatrix}$$

그러면 $A = a$ 는 $A^T A = 2$ 와 $A^T b = 2$ 를 만족한다. b 에서 b 의 A 위로의 사영 p 를 빼면 다음과 같다.

$$\text{첫 번째 단계} \quad B = b - \frac{A^T b}{A^T A} A = b - \frac{2}{2} A = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

이때 $A^T B = 0$ 이어야 한다. 이제 c 에서 c 의 A, B 위로의 사영을 빼서 C 를 얻는다. C 를 구하려면 사영을 두 번 한 뒤 빼야 한다.

$$\text{다음 단계} \quad C = c - \frac{A^T c}{A^T A} A - \frac{B^T c}{B^T B} B = c - \frac{6}{2} A + \frac{6}{6} B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

이때 $C = (1, 1, 1)$ 은 A, B 모두와 직교한다. 마지막으로 A, B, C 를 크기가 1인 벡터로 정규화한다. A, B, C 의 크기는 각각 $\sqrt{2}, \sqrt{6}, \sqrt{3}$ 이다. 이 크기로 각각 나누면 다음과 같이 정규직교 기저벡터 q_1, q_2, q_3 를 얻는다.

$$q_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad q_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \quad q_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

보통 A, B, C 에 분수가 포함된다. 또한 거의 항상 q_1, q_2, q_3 에 루트가 포함된다. 조금 복잡하지만 이는 $Q^T Q = I$ 를 얻는 데 따르는 대가이다.

$A = QR$ 분해

열이 a, b, c 인 행렬 A 로 시작해서 열이 q_1, q_2, q_3 인 행렬 Q 로 끝난다. 어떻게 두 행렬이 연관되는가? 벡터 a, b, c 는 벡터 q_1, q_2, q_3 의 선형결합이므로(그 역도 성립), 행렬 A 와 Q 를 연결하는 세 번째 행렬이 존재한다. 이 세 번째 행렬이 $A = QR$ 에 있는 삼각행렬 R 이다(이는 1장에서의 R 이 아니다).

첫 번째 단계는 $q_1 = \frac{a}{\|a\|}$ 이다(다른 벡터들과는 상관없다). 두 번째 단계는 식 (7)이다. 즉 벡터 b 는 A 와 B 의 선형결합이다. 여기서 C 와 q_3 는 상관없다.

다음과 같이 나중의 벡터가 이전 단계에 끼어들지 않는 점이 그람-슈미트 과정의 핵심이다.

- 벡터 a, A, q_1 은 모두 하나의 직선 위에 있다.
- 벡터 a, b, A, B, q_1, q_2 는 모두 하나의 평면 위에 있다.
- 벡터 $a, b, c, A, B, C, q_1, q_2, q_3$ 는 모두 하나의 부분공간(3차원) 위에 있다.

모든 단계에서 $a_1, \dots, a_k$ 는 $q_1, \dots, q_k$ 의 선형결합이다. 나중의 q_i 들은 상관없다. 상관행렬 R 은 삼각행렬이고 $A = QR$ 이다.

$$[a \ b \ c] = [q_1 \ q_2 \ q_3] \begin{bmatrix} q_1^T a & q_1^T b & q_1^T c \\ & q_2^T b & q_2^T c \\ & & q_3^T c \end{bmatrix} \Leftrightarrow A = QR \quad (9)$$

$A = QR$ 은 한 마디로 그람-슈미트 과정이다. Q^T 를 $A = QR$ 의 양변 왼쪽에 곱하면 식 (9)는 $R = Q^T A$ 이다.

(그람-슈미트 과정) 그람-슈미트 과정은 선형독립인 벡터 $a_1, \dots, a_n$ 을 정규직교벡터 $q_1, \dots, q_n$ 으로 만든다. 이들을 열로 가진 행렬은 $A = QR$ 을 만족한다. 나중의 q_i 들은 이전의 a_i 들과 직교하므로 $R = Q^T A$ 는 상삼각행렬이 된다.

앞서 살펴본 예에서 a_i 들과 q_i 들은 다음과 같다. $R = Q^T A$ 의 (i, j) 성분은 Q^T 의 i 행과 A 의 j 열의 내적이다. 내적 $q_i^T a_j$ 는 R 안에 들어간다. 그러면 다음과 같이 $A = QR$ 이 된다.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & -3 \\ 0 & -2 & 3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & -2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \sqrt{2} & \sqrt{18} \\ 0 & \sqrt{6} & -\sqrt{6} \\ 0 & 0 & \sqrt{3} \end{bmatrix} = QR$$

행렬 Q 와 R 을 자세히 살펴보자. A, B, C 의 크기는 각각 $\sqrt{2}, \sqrt{6}, \sqrt{3}$ 이며, 이는 R 의 주대각선 위에 있다. 행렬 Q 의 열들은 정규직교한다. 보통 QR 에는 루트 계산이 있어서, QR 이 LU 보다 복잡해 보인다. 두 분해 QR 과 LU 는 선형대수학에서 아주 중요하다.

선형독립인 열을 갖는 모든 행렬 A 는 $A = QR$ 분해를 갖는다. $m \times n$ 행렬 Q 는 정규직교하는 열을 갖는 행렬이고, 정사각행렬 R 은 양의 주대각성분을 갖는 상삼각행렬이다. 여기서 이 행렬들이 최소제곱 문제를 푸는 데 유용하다는 사실을 기억해두어야 한다. 이 사실을 자세히 살펴보자. 우선 행렬 A, Q, R 에 대해 다음이 성립한다.

$$A^T A = (QR)^T QR = R^T Q^T QR = R^T R$$

그러면 $A^T A \hat{x} = A^T b$ 는 $R^T R \hat{x} = R^T Q^T b$ 로 간단해지고, 결국 $R \hat{x} = Q^T b$ 가 된다(단, R 은 상삼각행렬).

$$\text{최소제곱 문제} \quad R^T R \hat{x} = R^T Q^T b \quad \text{또는} \quad R \hat{x} = Q^T b \quad \text{또는} \quad \hat{x} = R^{-1} Q^T b \quad (10)$$

해를 구할 수 없는 $Ax = b$ 를 푸는 대신, $R \hat{x} = Q^T b$ 를 역대입법으로 푸는 것이 훨씬 빠르다. 필요한 계산은 그람-슈미트 과정에 있는 mn^2 번의 곱셈이다. 이는 $A = QR$ 에서 직교하는 열을 갖는 행렬 Q 와 R 을 계산할 때 필요하다.

아래에는 그람-슈미트 과정에 대한 비공식적인 코드가 주어진다. 이는 식 (11)을 $j = 1, j = 2, \dots, j = n$ 인 경우에 대해 실행한다. 코드에서 중요한 4~5행에서는 $v = a_j$ 로부터 $v = a_j$ 의 각 q_i ($i < j$) 방향 위로의 사영을 뺀다. 코드의 마지막 행에서는 v 를 정규화($r_{jj} = \|v\|$ 로 나눔)하여 단위벡터 q_j 를 얻는다.

$$r_{kj} = \sum_{i=1}^m q_{ik} v_{ij}, \quad v_{ij} = v_{ij} - q_{ik} r_{kj}, \quad r_{jj} = \left(\sum_{i=1}^m v_{ij}^2 \right)^{1/2}, \quad q_{ij} = \frac{v_{ij}}{r_{jj}} \quad (11)$$

$a, b, c = a_1, a_2, a_3$ 로 시작하여, 이 코드는 다음과 같이 q_1 을 만든 후에 B, q_2 를 만든 후에 C , 그다음 q_3 를 만든다.

$$q_1 = \frac{a_1}{\|a_1\|}, \quad B = a_2 - (q_1^T a_2) q_1, \quad q_2 = \frac{B}{\|B\|}$$

$$C^* = a_3 - (q_1^T a_3) q_1, \quad C = C^* - (q_2^T C^*) q_2, \quad q_3 = \frac{C}{\|C\|}$$

식 (11)은 C^* 와 C 에서처럼 한 번에 하나의 사영만 뺀다. 이를 수정된 그람-슈미트(modified Gram-Schmidt) 과정이라 한다. 이 방법은 한꺼번에 모든 사영을 빼는 식 (8)보다 수치적으로 더 안정적이다. 아래 코드에서는 MATLAB 문법을 사용했다.

<pre> for j = 1:n v = A(:, j); for i = 1: j-1 R(i, j) = Q(:, i)'*v; v = v - R(i, j)*Q(:, i); end R(j, j) = norm(v); Q(:, j) = v/R(j, j); end </pre>	<pre> % n번 루프를 도는 수정된 그람-슈미트 과정 % v는 원래 행렬 A의 j열로 시작한다. % 열 q1부터 qj-1은 이미 행렬 Q 안에 있다. % Rij = qi^T aj를 계산한다. 이는 qi^T v와 같다. % 사영 (qi^T v)qi를 뺀다. % 이제 v는 모든 q1, ..., qj-1과 직교한다. % 행렬 R의 주대각성분 Rjj는 크기이다. % v를 자신의 크기로 나누어 다음의 qj를 얻는다. % for j = 1:n을 모두 진행하면 모든 qj를 얻는다. </pre>
---	---

행렬 A 에서 j 열을 복원하려면, 다음과 같이 코드의 마지막 단계와 중간 단계를 되돌려야 한다.

$$R(j, j)q_j = (v - (\text{자기 자신의 사영})) = (A \text{의 } j \text{열}) - \sum_{i=1}^{j-1} R(i, j)q_i \quad (12)$$

식 (12)에서 우변의 합에 대한 항을 좌변으로 이항하면, 식 (12)는 $QR = A$ 의 j 열이 된다.

고백하자면 MATLAB, 줄리아(Julia), 파이썬(Python)과 같은 훌륭한 시스템에서 사용하는 LAPACK 같은 좋은 소프트웨어는 그람-슈미트 코드를 사용하지 않을 수도 있다. 이제는 또 다른 방법이 있기 때문이다. '하우스홀더 반사(Householder reflections)'라는 반사변환이 행렬 A 에 작용하여 상삼각행렬 R 을 만든다. 이는 LU 분해에서 소거법이 상삼각행렬 U 를 만드는 것과 같은 방식으로, 한 번에 1개의 열씩 진행된다.

만약 A 가 삼중대각행렬이면 2×2 회전을 이용해 더 간단히 만들 수도 있다. 결과는 항상 $A = QR$ 이며 A 를 직교화하는 MATLAB 명령어는 $[Q, R] = \text{qr}(A)$ 이다. 그럼에도 불구하고 그람-슈미트 과정은 여전히 이해해두어야 할 좋은 방법이며, 가장 안정적인 코드는 각 단계에서 가장 큰 벡터 a_j 를 선택해 사용한다.

핵심 아이디어 리뷰

- ❶ 정규직교벡터 $q_1, \dots, q_n$ 이 Q 의 열이면 $q_i^T q_j = 0$, $q_i^T q_i = 1$ 이 성립한다. 이에 따라 $Q^T Q = I$ 가 성립한다.
- ❷ Q 가 정사각 직교행렬이면 $Q^T = Q^{-1}$ 이다. 즉 (전치행렬) = (역행렬)이다.
- ❸ Qx 의 크기는 x 의 크기와 같다($\|Qx\| = \|x\|$).
- ❹ q_i 들로 생성되는 행렬 Q 의 열공간 위로의 사영행렬은 $P = QQ^T$ 이다.
- ❺ Q 가 정사각행렬이면 $P = QQ^T = I$ 이고, 모든 벡터 b 는 $b = q_1(q_1^T b) + \dots + q_n(q_n^T b)$ 이다.
- ❻ 그람-슈미트 과정은 선형독립인 벡터 a, b, c 로부터 정규직교벡터 q_1, q_2, q_3 를 만든다. 이를 행렬로 표현하면 $A = (\text{직교행렬 } Q)(\text{삼각행렬 } R)$ 이다.

4.4 확인문제

4.4A 1 또는 -1로만 이루어진 두 열을 추가하여, 4×4 '아다마르 행렬(Hadamard matrix)' H_4 의 열들이 서로 직교하도록 만들어라. 어떻게 하면 H_4 를 직교행렬 Q_4 로 만들 수 있는가?

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad H_4 = \begin{bmatrix} 1 & 1 & x & x \\ 1 & -1 & x & x \\ 1 & 1 & x & x \\ 1 & -1 & x & x \end{bmatrix}, \quad Q_4 = \begin{bmatrix} & & & \\ & & & \\ & & & \\ & & & \end{bmatrix}$$

블록행렬 $H_8 = \begin{bmatrix} H_4 & H_4 \\ H_4 & -H_4 \end{bmatrix}$ 는 1과 -1로만 이루어진 H_4 다음의 아다마르 행렬이다. $H_8^T H_8$ 은 무엇인가?

벡터 $b = (6, 0, 0, 2)$ 를 H_4 의 1열 위로 사영한 것은 $p_1 = (2, 2, 2, 2)$ 이다. 또한 벡터 b 를 H_4 의 2열 위로 사영한 것은 $p_2 = (1, -1, 1, -1)$ 이다. 벡터 b 를 1열과 2열이 생성하는 공간 위로 사영한 $p_{1,2}$ 를 구하라.

풀이

H_4 는 다음과 같이 H_2 로부터 구할 수 있다. 이는 H_4 로부터 H_8 을 구한 것과 마찬가지이다.

$$H_4 = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \text{은 직교하는 열을 갖는다.}$$

그러면 $Q_4 = \frac{H_4}{2}$ 는 정규직교하는 열을 갖는다. 즉 2로 나누어 행렬 Q_4 의 열들의 크기가 1이 된다. 5×5 아다마르 행렬의 경우, 열들의 내적이 5개의 1 또는 -1의 합이므로 0이 될 수 없다. 따라서 5×5 아다마르 행렬은 만들 수 없다. H_8 은 크기가 $\sqrt{8}$ 인 직교하는 열들을 갖는다.

$$H_8^T H_8 = \begin{bmatrix} H_4^T & H_4^T \\ H_4^T & -H_4^T \end{bmatrix} \begin{bmatrix} H_4 & H_4 \\ H_4 & -H_4 \end{bmatrix} = \begin{bmatrix} 2H_4^T H_4 & 0 \\ 0 & 2H_4^T H_4 \end{bmatrix} = \begin{bmatrix} 8I & 0 \\ 0 & 8I \end{bmatrix} \cdot Q_8 = \frac{H_8}{\sqrt{8}}$$

4.4B 직교하는 열들에 대한 핵심은 무엇인가?

풀이

$Q^T Q$ 가 대각행렬이 되어 역행렬을 쉽게 구할 수 있다는 점이다. 이에 따라 직선 위로 사영한 뒤, 단지 더하기만 하면 된다. 축들은 서로 직교한다.

직교행렬 Q 와 직교사영 P

- A 는 n 개의 선형독립인 열을 $\mathbb{R}^m$ 안에서 가짐

$$m \times n \text{ 행렬 } A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$$

- $A^T A$ 는 대칭행렬이고 가역행렬
- P 는 A 의 열공간 위로의 사영행렬
- $P = A(A^T A)^{-1} A^T$
- $Ax = b$ 의 최소제곱해 : $A^T A \hat{x} = A^T b$

- Q 는 n 개의 정규직교하는 열을 $\mathbb{R}^m$ 안에서 가짐

$$Q = \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix}$$

- $Q^T Q = I$
- $Pb = (b \text{와 가장 가까운 점})$
- $P = QQ^T$
- $Qx = b$ 의 최소제곱해 : $\hat{x} = Q^T b$

- 그람-슈미트 직교화 과정 : 선형독립인 열 a_j , 정규직교하는 열 q_j

$$\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} = \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ 0 & 0 & \cdots & \\ 0 & 0 & 0 & r_{nn} \end{bmatrix}$$

- $A = QR$ 의 유사역행렬 A^+
: $A^+ = R^+ Q^+ = R^{-1} Q^T$

- (4.5절에서) $A = CR$ 의 유사역행렬 A^+

- Q 의 유사역행렬 Q^+
: $Q^+ = Q^T$

- $A^+ = R^+ C^+ = R^T (C^T A R^T)^{-1} C^T$

4.4 연습문제

※ [문제 1~12]는 직교하는 벡터와 직교행렬에 관한 문제이다.

1. 다음 벡터들이 서로 정규직교하는지, 아니면 직교만 하는지, 아니면 단지 선형독립인지 확인하라. 만약 정규직교벡터가 아니라면 두 번째 벡터를 바꿔서 정규직교벡터로 만들어라.

(a) $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

(b) $\begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix}, \begin{bmatrix} 0.4 \\ -0.3 \end{bmatrix}$

(c) $\begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}, \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}$

2. 두 벡터 $(2, 2, -1)$ 과 $(-1, 2, 2)$ 는 서로 직교한다. 각 벡터를 크기로 나눈 정규직교벡터 q_1, q_2 를 구하라. 그리고 q_1, q_2 를 행렬 Q 의 열로 구성하여 $Q^T Q, QQ^T$ 를 계산하라.

3. 다음 물음에 답하라.
 - (a) 행렬 A 는 크기가 모두 4인 직교하는 3개의 열을 갖는다. $A^T A$ 를 구하라.
 - (b) 행렬 A 는 크기가 각각 1, 2, 3인 직교하는 3개의 열을 갖는다. $A^T A$ 를 구하라.
4. 다음의 각 경우에 대한 예를 제시하라.
 - (a) 정규직교하는 열들을 갖지만 $QQ^T \neq I$ 를 만족하는 행렬 Q
 - (b) 서로 직교하지만 선형독립이 아닌 두 벡터
 - (c) 벡터 $q_1 = \frac{1}{\sqrt{3}}(1, 1, 1)$ 을 포함하는 $\mathbb{R}^3$ 의 정규직교기저
5. 평면 $x + y + 2z = 0$ 안에서 서로 직교하는 두 벡터를 찾으라. 그리고 이 두 벡터가 서로 정규직교하도록 만들어라.
6. Q_1 과 Q_2 가 직교행렬인 경우, $Q_1 Q_2$ 도 직교행렬임을 보여라($Q^T Q = I$ 를 사용하라).
7. Q 가 정규직교하는 열을 가질 때, $Qx = b$ 의 최소제곱해 $\hat{x}$ 는 무엇인가?
8. q_1, q_2 가 $\mathbb{R}^5$ 의 정규직교벡터일 때, 어떤 선형결합 $\underline{\hspace{1cm}} q_1 + \underline{\hspace{1cm}} q_2$ 가 임의로 주어진 벡터 b 와 가장 가까운가?
9. 다음 물음에 답하라.
 - (a) $q_1 = (0.8, 0.6, 0)$, $q_2 = (-0.6, 0.8, 0)$ 일 때, $P = QQ^T$ 를 계산하라. $P^2 = P$ 임을 확인하라.
 - (b) $Q^T Q = I$ 를 이용해 항상 $(QQ^T)^2 = QQ^T$ 가 성립함을 보여라. 그러면 $P = QQ^T$ 는 Q 의 열공간 위로의 사영행렬이다.
10. 정규직교벡터들은 자동으로 선형독립이다. 이 명제에 대한 다음 증명 과정에서 주어진 물음에 답하라.
 - (a) **[벡터 증명]** $c_1 q_1 + c_2 q_2 + c_3 q_3 = 0$ 이면 어떤 내적을 통해 $c_1 = 0$ 을 얻을 수 있는가? 비슷한 방법으로 $c_2 = 0$, $c_3 = 0$ 임을 알 수 있다. 따라서 벡터 q_i 들은 선형독립이다.
 - (b) **[행렬 증명]** $Qx = 0$ 이면 $x = 0$ 임을 보여라. 이때 Q 는 직사각행렬일 수 있으므로 Q^T 는 이용할 수 있지만 Q^{-1} 는 이용할 수 없다.
11. 다음 물음에 답하라.
 - (a) **[그람-슈미트 과정]** 두 벡터 $a = (1, 3, 4, 5, 7)$, $b = (-6, 6, 8, 0, 8)$ 로 생성되는 평면 위에서 정규직교벡터 q_1, q_2 를 찾으라.
 - (b) (a)의 평면 위에서 $(1, 0, 0, 0, 0)$ 과 가장 가까운 벡터를 찾으라.

12. a_1, a_2, a_3 가 $\mathbb{R}^3$ 의 직교하는 기저벡터일 때, 다음 식에 대해 $x_1 = \frac{a_1^T b}{a_1^T a_1}$ 임을 보여라.

$$b = x_1 a_1 + x_2 a_2 + x_3 a_3 \quad \text{또는} \quad \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = b$$

13. $b = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$ 로부터 $a = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ 에 어떤 수를 곱한 벡터를 뺄으로써 얻는 벡터를 B 라 하자. B 가 a 와 직교하도록 만들려면 a 에 어떤 수를 곱해야 하는가? a, b, B 를 나타낸 그림을 그려라.

14. $q_1 = \frac{a}{\|a\|}, q_2 = \frac{B}{\|B\|}$ 를 계산하고 다음과 같이 QR 분해를 함으로써, [문제 13]의 그람-슈미트 과정을 완성하라.

$$\begin{bmatrix} 1 & 4 \\ 1 & 0 \end{bmatrix} = [q_1 \ q_2] \begin{bmatrix} \|a\| & ? \\ 0 & \|B\| \end{bmatrix}$$

15. 다음 물음에 답하라.

(a) q_1, q_2 가 행렬 $A = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{bmatrix}$ 의 열공간을 생성하도록 하는 정규직교벡터 q_1, q_2, q_3 를 구하라.

(b) 4가지 기본 부분공간 중 q_3 를 포함하는 공간은 무엇인가?

(c) 최소제곱법을 이용해 $Ax = (1, 2, 7)$ 을 풀어라.

16. $a = (4, 5, 2, 2)$ 에 어떤 수를 곱해야 $b = (1, 2, 0, 0)$ 과 가장 가까운 벡터가 되는가? a, b 를 모두 포함한 평면 위에서 정규직교벡터 q_1, q_2 를 구하라.

17. 다음과 같은 벡터 a 방향 직선 위로의 벡터 b 의 사영 p 와 오차 $e = b - p$ 를 구하라. 그리고 정규직교 벡터 $q_1 = \frac{a}{\|a\|}, q_2 = \frac{e}{\|e\|}$ 를 구하라.

$$a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$$

18. (꼭 풀어볼 만한 문제) 다음과 같은 벡터 a, b, c 에 대해 그람-슈미트 과정을 적용하여 벡터 A, B, C 를 구하라.

$$a = (1, -1, 0, 0), \quad b = (0, 1, -1, 0), \quad c = (0, 0, 1, -1)$$

A, B, C 와 a, b, c 는 $d = (1, 1, 1, 1)$ 과 직교하는 벡터공간의 기저벡터이다.

19. $A = QR$ 이면 $A^T A = R^T R = (\text{삼각행렬}) \cdot (\text{삼각행렬})$ 이다. 행렬 A 에 그람-슈미트 과정을 적용하는 것은 $A^T A$ 에 소거법을 적용하는 것과 같다. $A^T A$ 의 피벗은 R 의 주대각성분의 제곱이 되어야 한다. 다음 행렬 A 에 그람-슈미트 과정을 적용하여 Q, R 을 찾으라.

$$A = \begin{bmatrix} -1 & 1 \\ 2 & 1 \\ 2 & 4 \end{bmatrix}, \quad A^T A = \begin{bmatrix} 9 & 9 \\ 9 & 18 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 9 & \\ & 9 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

20. 다음 명제의 참 또는 거짓을 판별하라. 그리고 각 경우에 대한 예를 제시하라.

(a) Q 가 직교행렬일 때 Q^{-1} 도 직교행렬이다.

(b) 3×2 행렬 Q 가 정규직교하는 열을 가지면, $\|Qx\|$ 는 항상 $\|x\|$ 와 같다.

21. 다음 행렬 A 의 열공간에 대한 정규직교기저를 구하고, 벡터 b 를 $C(A)$ 위로 사영하라.

$$A = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 1 & 1 \\ 1 & 3 \end{bmatrix}, \quad b = \begin{bmatrix} -4 \\ -3 \\ 3 \\ 0 \end{bmatrix}$$

22. 벡터 $a = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $c = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}$ 에 대해 그람-슈미트 과정을 적용하여 서로 직교하는 벡터 A, B, C 를 구하라.

23. 정규직교하는 열 q_1, q_2, q_3 를 선형독립인 열 a, b, c 의 선형결합으로 표현하라. 그다음 행렬 A 를 QR 로 분해하라.

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix}$$

24. 다음 물음에 답하라.

(a) $x_1 + x_2 + x_3 - x_4 = 0$ 의 모든 해로 생성되는 $\mathbb{R}^4$ 의 부분공간 S 의 기저를 구하라.

(b) 직교여공간 $S^\perp$ 의 기저를 구하라.

(c) $b_1 + b_2 = b = (1, 1, 1, 1)$ 이 되도록 하는 벡터 b_1 을 S 에서, 벡터 b_2 를 $S^\perp$ 에서 찾으라.

25. $ad - bc > 0$ 이면 $A = QR$ 의 성분은 다음과 같다.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{\begin{bmatrix} a & -c \\ c & a \end{bmatrix}}{\sqrt{a^2 + c^2}} \frac{\begin{bmatrix} a^2 + c^2 & ab + cd \\ 0 & ad - bc \end{bmatrix}}{\sqrt{a^2 + c^2}}$$

$a, b, c, d = 2, 1, 1, 1$ 일 때와 $a, b, c, d = 1, 1, 1, 1$ 일 때의 $A = QR$ 을 각각 작성하라. 열들이 선형종속이어서 그람-슈미트 과정이 실패할 때, R 의 어떤 성분이 0이 되는가?

※ [문제 26~29]에서는 식 (11)의 QR 분해에 대한 코드를 사용한다. 이 코드는 그람-슈미트 과정을 구현한다.

26. (식 (11) 이후의 단계에서 C^* 를 통해 발견된) C 가 식 (8)의 C 와 같은 이유를 설명하라.

27. 식 (8)에서는 c 로부터 c 의 A, B 방향의 성분을 뺀다. 왜 a, b 방향의 성분을 빼지는 않는가?

28. 식 (11)의 어떤 부분에서 mn^2 번의 곱셈이 필요한가?

29. 세 벡터 $a = (2, 2, -1)$, $b = (0, -3, 3)$, $c = (1, 0, 0)$ 에 MATLAB QR 코드를 적용하여 q_i 들을 구하라.

※ [문제 30~35]에서는 특별한 직교행렬들을 다룬다.

30. 처음 4개의 웨이블릿(wavelet)이 다음 웨이블릿 행렬(wavelet matrix) W 의 열에 있다.

$$W = \frac{1}{2} \begin{bmatrix} 1 & 1 & \sqrt{2} & 0 \\ 1 & 1 & -\sqrt{2} & 0 \\ 1 & -1 & 0 & \sqrt{2} \\ 1 & -1 & 0 & -\sqrt{2} \end{bmatrix}$$

이 행렬의 열들의 특징은 무엇인가? 웨이블릿 행렬의 역행렬 W^{-1} 를 구하라.

31. 다음 행렬 Q 가 직교행렬이 되도록 하는 c 를 구하라.

$$Q = c \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$$

그리고 벡터 $b = (1, 1, 1, 1)$ 을 행렬 Q 의 1열 위로 사영하라. 그다음 벡터 b 를 처음 2개의 열이 있는 평면 위로 사영하라.

32. u 가 단위벡터일 때 $Q = I - 2uu^T$ 는 반사행렬이다([EX 3] 참고). 벡터 $u = (0, 1)$ 로부터 행렬 Q_1 을 찾고, 벡터 $u = \left(0, \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ 로부터 행렬 Q_2 를 찾으라. 또한 Q_1, Q_2 가 각각 벡터 $(1, 2), (1, 1, 1)$ 에 곱해질 때의 반사를 그림으로 나타내라.

33. 직교행렬인 동시에 하삼각행렬인 모든 경우를 찾으라.

34. $Q = I - 2uu^T$ 는 $u^T u = 1$ 일 때 반사행렬이다. 두 번 반사하면 $Q^2 = I$ 가 된다. 다음 물음에 답하라.

(a) $Qu = -u$ 임을 증명하라. 벡터 u 는 반사면인 거울과 직교한다.

(b) $u^T v = 0$ 일 때 Qv 를 구하라. 반사면은 벡터 v 를 포함한다. 이때 벡터 v 는 자기 자신으로 반사된다.

도전문제

35. (MATLAB 문제) $A = \text{eye}(4) - \text{diag}([1 \ 1 \ 1], -1)$ 일 때, $[Q, R] = \text{qr}(A)$ 를 구하라. 이를 통해 행렬 A 의 열 $(1, -1, 0, 0)$, $(0, 1, -1, 0)$, $(0, 0, 1, -1)$, $(0, 0, 0, 1)$ 을 직교화한다. 행렬 Q 의 열들에 적당한 수를 곱하여, 행렬 Q 의 성분을 모두 정수로 만들 수 있는가?

36. $m \times n$ 행렬 A 의 랭크가 n 이다. $\text{qr}(A)$ 는 정사각행렬 Q 와 R 아래쪽에 0 을 만든다.

$$\text{MATLAB으로부터 인수는 } (m \times m \text{ 행렬}) \cdot (m \times n \text{ 행렬}) \quad A = [Q_1 \ Q_2] \begin{bmatrix} R \\ 0 \end{bmatrix}$$

Q_1 의 n 개의 열은 어떤 기본 부분공간의 기저를 이루는가? 또한 Q_2 의 $(m-n)$ 개의 열은 어떤 기본 부분공간의 기저를 이루는가?

37. $m \times n$ 행렬 Q 의 열공간 위로의 사영행렬은 $P = QQ^T$ 이다. 하나의 열 a 를 추가하여 행렬 $A = [Q \ a]$ 를 만들었다고 하자. 그람-슈미트 과정은 벡터 a 를 어떤 벡터 q 로 바꾸는가? 이를 위해 a 로 시작하여 _____을(를) 빼고, _____(으)로 나누어 q 를 구하라.

4.5 행렬의 유사역행렬

랭크 $r = m = n = 2$

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$I = (I \text{의 역행렬}) \\ (I)(I) = I$$

랭크 $r = m < n = 3$

$$I_L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$I_L = (I_R \text{의 좌역행렬}) \\ (I_L)(I_R) = I$$

랭크 $r = n < m = 3$

$$I_R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$I_R = (I_L \text{의 우역행렬}) \\ (I_R)(I_L) \neq I$$

첫 번째 행렬 I 만 반드시 가역행렬이다. I_L , I_R 은 한쪽 방향 역행렬만 가지며 양쪽 방향 역행렬

을 갖지 않는다. 모든 행렬 A 는 유사역행렬 A^+ 를 갖는다. 유사역행렬은 역행렬, 좌역행렬, 우역행렬을 모두 포함한다. 우선 한쪽 방향 역행렬의 특징에 대해 정리하면 다음과 같다.

- | | |
|--|---|
| <ul style="list-style-type: none"> • A가 좌역행렬을 가짐 : $A^+A = I$ • A가 선형독립인 열을 가짐 • A는 길고 좁은 행렬 : 랭크 $r = n$ • $Ax = b$는 해가 존재하지 않거나 유일한 해를 가짐 • 영공간 $N(A) = \{ \text{영벡터} \}$ • $A^T A$는 $n \times n$ 가역행렬 • 좌역행렬은 $A^+ = (A^T A)^{-1} A^T$ | <ul style="list-style-type: none"> • A가 우역행렬을 가짐 : $AA^+ = I$ • A가 선형독립인 행을 가짐 • A는 짧고 넓은 행렬 : 랭크 $r = m$ • $Ax = b$는 유일한 해 혹은 무수히 많은 해를 가짐 • 좌영공간 $N(A^T) = \{ \text{영벡터} \}$ • AA^T는 $m \times m$ 가역행렬 • 우역행렬은 $A^+ = A^T (AA^T)^{-1}$ |
|--|---|

위 표에서 $A^+A = I$ 에 대한 왼쪽 부분을 살펴보자. 랭크는 $r = n < m$ 이고, m 개의 방정식 $Ax = b$ 는 방정식이 너무 많아서 해가 존재하지 않을 수도 있다. 그러면 $\hat{x} = A^+b$ 는 $A^T A \hat{x} = A^T b$ 를 푸는 해가 되며, 오차벡터 $e = b - A\hat{x}$ 는 최소가 된다. 이것이 바로 최소제곱법이다.

이번에는 위 표에서 $AA^+ = I$ 에 대한 오른쪽 부분을 살펴보자. 랭크는 $r = m < n$ 이고, $Ax = b$ 는 무수히 많은 해를 갖는다. 이 경우 $x^+ = A^+b$ 는 $Ax = b$ 의 최소길이 해(최단거리 해)이다. 이 해는 행렬 A 의 행공간 안에 있고 다른 해들은 $x = x^+ + x_n$ 으로 표현된다. 여기서 영공간의 성분 x_n 은 해의 길이를 증가시킨다.

행렬 A 의 랭크 r 이 $r < m$ 과 $r < n$ 을 만족하면 무엇을 기대할 수 있는가? 이 경우에는 한쪽 방향 역행렬도 없다. 유사역행렬 A^+ 가 최적의 $x^+ = A^+b$ 를 제공한다. 벡터 x^+ 는 오차 $\|Ax - b\|^2$ 을 최소화한다. 동일한 Ax 값과 동일한 오차를 제공하는 벡터 $x = x^+ + x_n$ 들 중 벡터 x^+ 가 최소길이를 갖는다. 정리하면 다음과 같다.

x^+ 는 $Ax = b$ 의 '최소노름 최소제곱해'가 된다.

$\|Ax - b\|^2$ 을 최소화하는 모든 벡터들 중 $x^+ = A^+b$ 가 $\|x\|^2$ 을 최소화한다.

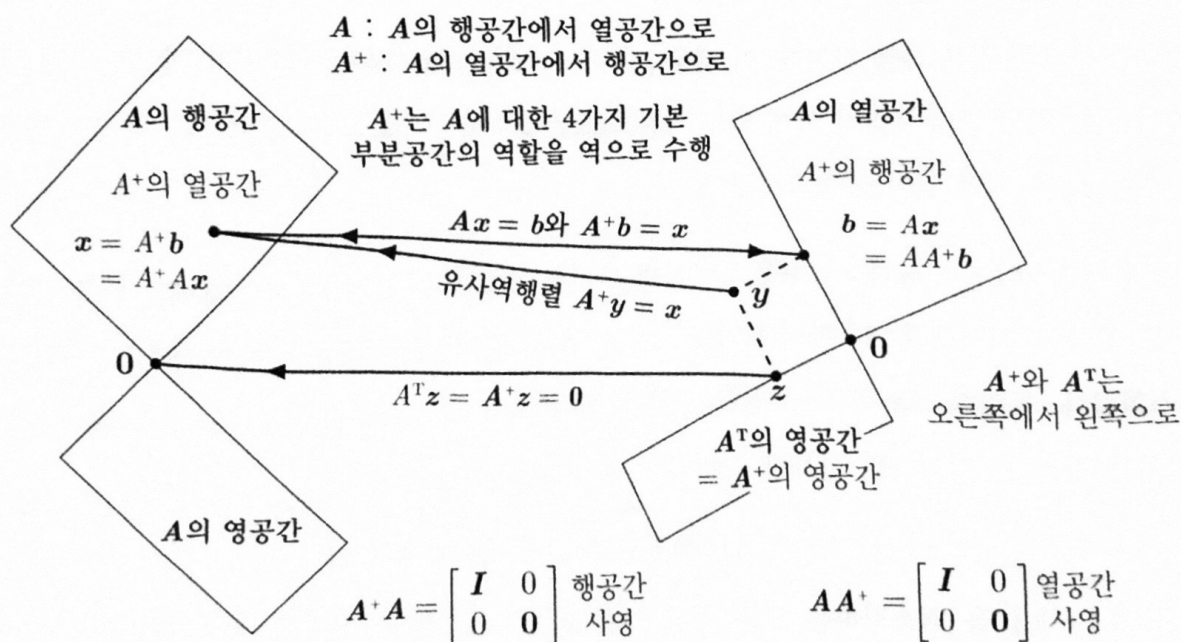
여기서 핵심은 x^+ 가 영공간 성분으로 0을 갖는다는 것이다. 만약 행렬 A 가 선형종속인 열을 갖는다면 A^+ 가 필요하다. 이 절은 필수로 알아두어야 할 내용은 아니지만, 이와 같은 이유로 4장의 절로 포함했다. 4가지 기본 부분공간을 이용해 A^+ 를 설명할 것이다. A^+ 는 $A = CR$, $A = QR$ 또는 특잇값 분해(SVD)를 이용해 계산할 수 있다.

$m \times n$ 행렬 A 의 $n \times m$ 유사역행렬 A^+

- 모든 m 차원 벡터 y 는 직교하는 두 부분으로 나누어 $y = b + z$ 로 표현한다.
- 벡터 b 는 A 의 열공간 안에 있고, 벡터 z 는 A^T 의 영공간 안에 있다. $A^T z = 0$ 이고 $A^+ z = 0$ 이다.
- 행렬 A 의 행공간 안의 어떤 벡터 x 에 대해 $Ax = b$ 이다. $A^+ b = x$ 를 찾기 위해 $Ax = b$ 에 역변환을 취한다.

A 의 유사역행렬은 $A^+ b = x$, $A^+ z = 0$, $A^+ y = A^+ b + A^+ z = x$ 를 만족한다.

이제 A^+ 에 관한 큰 그림을 살펴보자([그림 4.11] 참고). A^+ 는 A^T 와 동일한 4가지 기본 부분 공간을 갖는다. 차이점은 $A^+ A$ 의 경우 행공간 안에 있는 벡터 x 를 다시 동일한 벡터 x 로 되돌려놓는다는 것이다. 사실 $A^+ A$ 는 행렬 A 의 행공간 위로 사영하는 $n \times n$ 사영행렬 P_{row} 이다. 대칭적으로 AA^+ 는 행렬 A 의 열공간 위로 사영하는 $m \times m$ 사영행렬 P_{column} 이다. 가능한 경우 A^+ 는 A 를 역변환한다(열공간에서 행공간으로).



[그림 4.11] A 의 열공간 안에 있는 $b = Ax$ 는 A 의 행공간 안에 있는 $x = A^+b$ 로 돌아간다.

$A^+ z = 0$ 이므로 임의의 $y = b + z$ 도 $x = A^+ b$ 로 돌아간다. 그러면 $(A^+ A)^2 = A^+ A$ 이다.

A^+ 는 무어(Moore)가 처음 제안했고, 2020년 노벨 물리학상 수상자인 펜로즈(Penrose)가 다시 제안했다. 다음과 같이 A^+ 는 대칭인 사영행렬 P_{row} 와 P_{column} 을 만들 때 사용된다.

$$P_{\text{row}} = A^+ A = (A^+ A)^2 = (A^+ A)^T, \quad P_{\text{column}} = A A^+ = (A A^+)^2 = (A A^+)^T$$

$A = CR$ 또는 $A = QR$ 로부터 A^+ 를 계산할 때, A^+ 가 만족해야 하는 특징에 대한 판정법은 유용하다.

EX 1 다음과 같은 행렬 A , 벡터 y , 유사역행렬 A^+ 를 살펴보자.

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ y_2 \end{bmatrix} = b + z, \quad A^+ = \begin{bmatrix} 1/2 & 0 \\ 0 & 0 \end{bmatrix}$$

벡터 $b = \begin{bmatrix} y_1 \\ 0 \end{bmatrix}$ 은 행렬 A 의 열공간 안에 있다.

$$Ax = b \Leftrightarrow \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1/2 \\ 0 \end{bmatrix} = \begin{bmatrix} y_1 \\ 0 \end{bmatrix}$$

벡터 $z = \begin{bmatrix} 0 \\ y_2 \end{bmatrix}$ 는 행렬 A^T 의 영공간 안에 있다.

$$A^T z = 0 \Leftrightarrow \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

따라서 $A^+ y = A^+ b + A^+ z = x + 0 = \begin{bmatrix} y_1/2 \\ 0 \end{bmatrix}$ 이다. 그러면 $A^+ = \begin{bmatrix} 1/2 & 0 \\ 0 & 0 \end{bmatrix}$ 이고,

$A^+ A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ 이다.

위 결과는 다음과 같이 아주 쉽고 유용한 경우인 대각행렬 D (정사각행렬 혹은 직사각행렬)로 확장된다.

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{의 유사역행렬은 } D^+ = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{이다.}$$

위 식에서는 가능한 경우에만 역수를 취했다. 그러나 $1/0$ 은 0 으로 표기했다. 대각행렬 D 가 추가로 영행(또는 영열)을 갖는다면 D^+ 는 추가로 영열(또는 영행)을 갖는다.

이 예를 통해 B , C 가 가역행렬일 때, 대각행렬 D 를 임의의 행렬 $A = BDC$ 로 확장할 수 있다.

A^+ 는 $C^{-1} D^+ B^{-1}$ 이다. 여기서 B , C 의 역행렬을 취할 수 있지만, D 안에 있는 0 의 역수를

사용할 수는 없다. 그러므로 D^+ 는 D^{-1} 를 대치하고 A^+ 는 A^{-1} 를 대치한다. 예를 들어 다음과 같다.

• $A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ 의 유사역행렬은 A 의 좌역행렬 $A^+ = \frac{1}{2} \begin{bmatrix} 1 & 1 \end{bmatrix}$ 이다.

• $A = \begin{bmatrix} 1 & 1 \end{bmatrix}$ 의 유사역행렬은 A 의 우역행렬 $A^+ = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ 이다.

대부분의 행렬이 $A^{-1}A = I$ 와 $AA^{-1} = I$ 를 만족하는 역행렬 A^{-1} 를 갖지 못한다. 가역행렬은 선형독립인 행과 열을 갖는 정사각행렬뿐이다. A^{-1} 가 존재하려면 A 의 영공간과 A^T 의 영공간 안에 영벡터만 존재해야 한다.

이 장에서는 선형독립인 열을 가진 길고 좁은 행렬($m > n$)에 대해 살펴봤다. 이러한 행렬 A 의 랭크는 $r = n$ 이고, 영공간 $N(A)$ 안에는 영벡터만 존재한다. 그러면 $n \times n$ 행렬 $A^T A$ 는 가역행렬이다. 이 경우 $A^+ = (A^T A)^{-1} A^T$ 는 A 의 좌역행렬이다.

$$A^+ A = (A^T A)^{-1} A^T A = I \text{이다. 그러나 } AA^+ = I \text{는 아니다.} \quad (1)$$

다른 가능성은 선형독립인 행을 가진 짧고 넓은 행렬($m < n$)이다. 이러한 행렬 A 의 랭크는 $r = m$ 이고, $N(A^T)$ 안에는 영벡터만 존재한다. 그러면 $m \times m$ 행렬 AA^T 는 가역행렬이다. 이 경우 $A^+ = A^T (AA^T)^{-1}$ 는 A 의 우역행렬이다.

$$AA^+ = AA^T (AA^T)^{-1} = I \text{이다. 그러나 } A^+ A = I \text{는 아니다.} \quad (2)$$

행렬 A 의 중요한 역할 : 행공간에서 열공간으로

행렬 A 의 행공간과 열공간에 관한 중요한 사실을 알아보자. 우선 행공간과 열공간의 차원 $r(A)$ (랭크)이 동일하다는 것은 이미 잘 알려진 사실이다. 그러나 그것이 전부는 아니다. 더 중요한 점은 행공간의 두 벡터 x 와 X 가 열공간의 동일한 벡터 $Ax = AX$ 로 갈 수는 없다는 것이다. 열공간의 모든 벡터 $b = Ax$ 는 행공간의 단 하나의 벡터 x 로부터 나온다. 그러면 $A^+ b = x$ 이다.

이 사실은 쉽게 증명할 수 있다. 우선 x 와 X 는 모두 행렬 A 의 행공간 안에 있고 $Ax = b$, $AX = b$ 라 가정하자. $A(x - X) = b - b = 0$ 이므로 $x - X$ 는 행렬 A 의 영공간 안에 있어야 한다. 그러나 $x - X$ 는 행공간 안에 있기도 하다(벡터공간은 뺄셈에 대해 닫혀있으므로). 영공간은 항상 행공간과 직교하므로, 두 공간 안에 동시에 존재하는 $x - X$ 는 자기 자신과 직교해야 한다. 그러므로 $x - X$ 는 반드시 영벡터여야 한다.

이 아이디어를 다른 방향으로 설명할 수 있다. $x_1, \dots, x_r$ 이 행렬 A 의 행공간의 기저벡터라고 하자. 그러면 $Ax_1, \dots, Ax_r$ 은 행렬 A 의 열공간의 기저벡터가 된다. 이 결과를 통해 행공간과 열공간의 차원이 모두 r 이 된다는 사실을 쉽게 이해할 수 있다(1장의 중요한 사실). 1장에서 배우고 3장에서 소거법을 통해 완성한 $A = CR$ 분해가 여기에 사용된다. $A = CR$ 은 유사역행렬 A^+ 에 대해 다음과 같은 사실을 알려준다.

$m \times r$ 행렬 C 는 행렬 A 로부터 가져온 r 개의 선형독립인 열들로 구성된다. $r \times n$ 행렬 R 은 행렬 A 의 기약행 사다리꼴로부터 가져온 r 개의 선형독립인 행들로 구성된다. 이 두 행렬을 통해 $A = CR$ 분해를 얻는다. 이제 $A = CR$ 의 유사역행렬은 $A^+ = R^+ C^+$ 가 된다.

$A = CR$ 의 유사역행렬 $A^+ = R^+ C^+$

여기서는 C 와 R 의 한쪽 방향 역행렬을 사용한다. C 는 r 개의 선형독립인 열을 갖고, R 은 r 개의 선형독립인 행을 갖는다. $C^T C$ 와 RR^T 는 $r \times r$ 가역행렬이다. C 와 R 의 유사역행렬은 한쪽 방향 역행렬이다($C^+ C = I$, $RR^+ = I$).

$$C^+ = (C^T C)^{-1} C^T, \quad R^+ = R^T (RR^T)^{-1} \quad (3)$$

운 좋게도 C 의 좌역행렬과 R 의 우역행렬이 $R^+ C^+$ 의 순서로 배치되어 $A = CR$ 의 유사역행렬에 대해 알려준다. 이 사실은 중요하다([연습문제 10] 참고).

유사역행렬

$$\begin{aligned} A^+ &= R^+ C^+ = R^T (RR^T)^{-1} (C^T C)^{-1} C^T \\ &= R^T (C^T C R R^T)^{-1} C^T = R^T (C^T A R^T)^{-1} C^T \end{aligned} \quad (4)$$

이렇게 잘 맞아떨어지는 여러 사실에 대해 증명이 필요하다고 생각할 수 있다. 핵심은 C 가 선형독립인 열을 갖고 R 이 선형독립인 행을 가질 때, $C^T C$ 와 RR^T 가 가역행렬이라는 것이다. 이 절에서는 A^+ 에 대한 식 (4)를 적용하는 예를 소개한다. 이 주제는 이 책의 웹사이트 math.mit.edu/linearalgebra에서 이어서 소개한다.

지금까지 A^+ 를 A 의 4가지 기본 부분공간으로 설명했다. 그러나 무어와 펜로즈는 다른 접근 방식을 취했다. 펜로즈는 유사역행렬 A^+ 에 대해 다음 4가지 요구 조건을 만들었으며, 이 조건들을 만족하는 유일한 행렬이 바로 유사역행렬이다.

$$AA^+A = A, \quad A^+AA^+ = A^+, \quad (A^+A)^T = A^+A, \quad (AA^+)^T = AA^+ \quad (5)$$

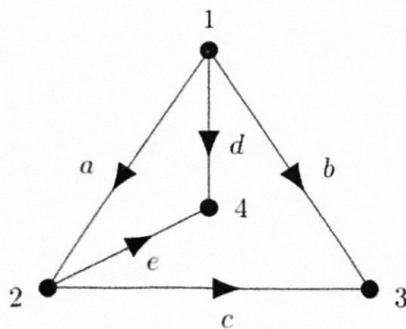
그러면 A^+A 와 AA^+ 는 각각 행렬 A 의 행공간과 열공간 위로의 사영행렬이다.

좋은 성질을 갖는 기저를 찾으려면 많은 노력이 필요하다. 그중에서도 매우 뛰어난 성질은 '직교성'이다. 7장의 핵심 주제는 행렬 A 의 특이벡터(singular vector)를 찾는 것이다. 특이벡터는 행렬 A 의 행공간 안에 있는 정규직교벡터 $v_1, \dots, v_r$ 이고, 각 벡터에 행렬 A 를 곱한 $Av_1, \dots, Av_r$ 은 행렬 A 의 열공간 안에서 서로 직교한다. 응용 선형대수학에서 이러한 구조는 그야말로 완벽하다고 말할 수 있다.

유사역행렬 A^+ 는 $A = CR$, $A = QR$, 특잇값 분해(SVD) $A = U\Sigma V^T$ 와 같은 다양한 분해를 통해 동일하게 얻을 수 있다.

A 로부터 A^+ 를 구하는 예 : 그래프의 결합행렬

3.5절의 $m \times n$ 결합행렬은 그래프의 모든 에지에 대해 하나의 행을 갖고, 모든 노드에 대해 하나의 열을 갖는다. i 행에는 0이 아닌 두 성분 -1 , 1 이 있다. -1 은 에지 i 의 출발노드를 나타내고, 1 은 에지 i 의 도착노드를 나타낸다. 실제 흐름은 화살표 방향을 따를 수도 있고, 그렇지 않을 수도 있다. 단지 각 에지에 대해 어떤 방향이 양의 방향인지에 대한 기준만 잡아놓은 상태이다.



$$A = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 \end{bmatrix} \quad \begin{array}{l} \text{에지} \\ a \\ b \\ c \\ d \\ e \end{array}$$

$$A \text{의 (4열)} = -\{ (1\text{열}) + (2\text{열}) + (3\text{열}) \} \\ \text{따라서 } R \text{의 (4열)} = -1, -1, -1$$

행렬 A 에서 $m = 5$, $n = 4$ 이고 랭크는 $r = 3$ 이다. 처음 세 열은 다음과 같이 선형독립이다.

$$A = CR = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \quad (6)$$

선형방정식 $Ax = b$ 에 대한 기본적인 사실들을 다시 생각해보자. 벡터 b 가 행렬 A 의 열들의 결합일 때 $Ax = b$ 의 해가 존재한다. 최소노름해 x_0 는 행렬 A 의 행공간 안에 있는 유일한 해이다 (즉 x_0 는 행들의 선형결합이다). 다른 해 $x = x_0 + x_n$ 은 더 큰 노름 $\|x\|^2 = \|x_0\|^2 + \|x_n\|^2$ 을 갖는다. $x = x_0 + x_n$ 은 $Ax_n = 0$ 을 만족하는 영공간에 포함된다.

모든 행렬 A 는 행공간에서 열공간으로(즉 r 차원에서 r 차원으로)의 가역변환이다.

$$\begin{array}{lcl} x_0 \text{는 } A \text{의 행공간 안에 있음} & \xrightarrow{Ax_0 = b_0} & A^+ b_0 = x_0 \quad b_0 \text{는 } A \text{의 열공간 안에 있음} \\ x_n \text{는 } A \text{의 영공간 안에 있음} & \xrightarrow{Ax_n = 0} & A^+ b_n = 0 \quad b_n \text{는 } A^T \text{의 영공간 안에 있음} \end{array}$$

A^+ 는 어떻게 계산해야 할까? 추천하는 공식은 **특잇값 분해**(SVD; Singular Value Decomposition) $A = U \Sigma V^T$ 로부터 나오는 $A^+ = V \Sigma^+ U^T$ 이다. 여기서 U 와 V 는 $U^{-1} = U^T = U^+$, $V^{-1} = V^T = V^+$ 를 만족하는 정사각 직교행렬이다. $m \times n$ 대각행렬 Σ 는 $n \times m$ 대각 유사역행렬 Σ^+ 를 갖는다(Σ 의 각 $\sigma_i > 0$ 은 Σ^+ 의 $1/\sigma_i$ 로 대체된다). Σ 의 주대각선에 0이 있으면 Σ^+ 의 주대각선에서도 0으로 놓는다.

만약 A 가 결합행렬처럼 정수로만 이루어진 작은 행렬이면 A^+ 를 계산하는 또 다른 방법이 있다. 식 (6)에서는 SVD 대신 $A = CR$ 로 표현했다. 그러면 식 (4)에 따라 $A^+ = R^+ C^+$ 로 계산할 수 있다. 이 방식에서는 유리수는 포함되지만 무리수는 포함되지 않는다(SVD에서는 무리수가 포함된다).

$$A = CR \text{의 유사역행렬} \quad A^+ = R^T (C^T A R^T)^{-1} C^T \quad (7)$$

A^+ 와 A^T 는 동일한 행공간과 열공간을 갖는다. 둘의 차이점은 $A^+ A$ 가 행공간 위로 사영하는 대칭인 사영행렬이라는 것이다. 식 (5)의 4가지 펜로즈 조건을 확인할 수 있고, 다음과 같이 식 (7)로부터 $(A^+ A)^2 = A^+ A$ 임을 확인할 수 있다.

$$(A^+ A)^2 = R^T (C^T A R^T)^{-1} C^T A R^T (C^T A R^T)^{-1} C^T A = A^+ A \quad (8)$$

비슷하게 $(A A^+)^2 = A A^+$ 이다. 만약 A 가 가역행렬이면 $A^+ = A^{-1}$ 이다. 앞서 살펴본 그래프의 결합행렬의 경우, 다음과 같이 식 (7)을 이용한 A^+ 의 계산은 정확한 결과를 낸다.

$$C^T A R^T = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ -1 & -1 & 2 \end{bmatrix}, \quad A^+ = \frac{1}{8} \begin{bmatrix} -2 & -2 & 0 & -2 & 0 \\ 2 & 0 & -2 & 0 & -2 \\ 0 & 3 & 3 & -1 & -1 \\ 0 & -1 & -1 & 3 & 3 \end{bmatrix} \quad (9)$$

A^T 의 4가지 기본 부분공간은 A^+ 의 4가지 기본 부분공간과 동일하다. A^T 와 A^+ 의 영공간을 확인하라. 행공간과 열공간은 직교성에 의해 반드시 일치해야 한다.

A 의 행들을 더하면 0이다. $\rightarrow A^+$ 의 열들을 더하면 0이다.

A 의 (2행) = (1행) + (3행) $\rightarrow A^+$ 의 (2열) = (1열) + (3열)

A 의 (1행) = (4행) - (5행) $\rightarrow A^+$ 의 (1열) = (4열) - (5열)

방정식 $A^T y = 0$ 은 키르히호프의 전류법칙이다(각 노드에서 들어오는 전류와 나가는 전류가 같다). 이 식은 회로 이론과 계산 공학에서 매우 중요하다.

4.5 연습문제

1. 다음 행렬에 대해 유사역행렬 A^+ 를 구하라.

$$A_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2 & 0 \\ 0 & 4 \\ 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \end{bmatrix}$$

2. 행렬 A 의 랭크가 1이면 $A = cr^T = (\text{하나의 열과 하나의 행의 곱})$ 으로 나타낼 수 있다. $A^+ = rc^T / (r^T r)(c^T c)$ 가 식 (5)의 펜로즈 조건을 만족함을 보여라.

3. 행렬 A 에 영행(0으로만 이루어진 행)을 추가하면 A^+ 에 영열(0으로만 이루어진 열)이 추가됨을 보여라.

힌트 A 에 영행이 추가되고 A^+ 에 영열이 추가된 경우, 4가지 펜로즈 조건을 계속 만족하는지 확인한다.

4. 다음 행렬 A 의 행공간 $C(A^T)$ 와 열공간 $C(A)$ 는 무엇인가? A^T 와 동일한 4가지 기본 부분공간을 갖는 유사역행렬 A^+ 를 구하라.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

5. $A = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \end{bmatrix}$ 과 A^T 의 유사역행렬을 구하라. 그리고 $(A^T)^+ = (A^+)^T$ 에 대해 설명하라.

6. P 가 사영행렬일 때($P^T = P = P^2$), $P^+ = P$ 가 식 (5)의 펜로즈 조건을 만족함을 보여라.
7. 임의의 행렬 $A = CR$ 에 대해 식 (8)을 통해 $(A^+A)^2 = A^+A$ 가 성립함을 확인했다. 비슷하게 $(AA^+)^2 = AA^+$ 가 성립함을 확인하라.
8. 임의의 행렬 A 에 대해 $A^+ = R^T(C^T A R^T)^{-1} C^T$ 와 $A = CR$ 을 곱해서 A^+A 와 AA^+ 가 대칭행렬임을 보여라. 이 결과를 [문제 7]의 $(A^+A)^2 = A^+A$ 와 결합하면 A^+A 와 AA^+ 가 대칭인 사영행렬임을 확인할 수 있다.
9. 274쪽의 그래프에서 노드 4와 노드 4로부터 나가는 2개의 에지를 지우자(즉 삼각형 그래프는 그대로 둔다). 결합행렬 A 를 작성하고 A^+ 를 구하라. 그리고 $A = CR$ 형태로 분해한 뒤, 식 (7)을 이용해 A^+ 를 다시 계산하라.
10. $A = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ 일 때, $(AB)^+$ 는 B^+A^+ 가 아니고 $(BA)^+ = A^+B^+$ 임을 보여라. 이때 첫 번째 인수는 선형독립인 열이 필요하고 두 번째 인수는 선형독립인 행이 필요하다.
11. 왜 A^+ 의 유사역행렬은 항상 A 와 동일한가?

4장에 대한 고찰 : 직교성

이 책에서 다룬 주제를 되돌아보면, 직교벡터와 직교행렬이 가장 중요하다고 할 수 있다. 이 중요한 주제는 1장에서 다룬 벡터의 크기와 내적 개념까지 전파된다. 이와 관련한 몇 가지 사실을 살펴보자.

- ❶ Qx 의 크기는 x 의 크기와 같다. : $(Qx)^T(Qx) = x^T Q^T Qx = x^T x = \|x\|^2$
- ❷ 내적 $(Qx)^T(Qy)$ 는 내적 $x^T y$ 와 같다. : $x^T Q^T Qy = x^T y$
- ❸ $m \times n$ 행렬 Q 의 열공간 위로의 사영행렬은 $P = QQ^T = (QQ^T)^2$ 이다.
- ❹ ($m > n$ 일 때) $Qx = b$ 의 최소제곱해는 $\hat{x} = Q^T b = Q^+ b$ (유사역행렬)이다.

만약 Q 가 정사각행렬이면 $Q^T = Q^{-1}$ 이다. 더 자세한 내용은 5, 6, 7장에서 다룬다.

- ⑤ 거듭제곱 Q^N 과 곱 $Q_1 Q_2 \cdots Q_N$ 은 직교행렬이다. : $|(행렬식)| = 1$
- ⑥ 대칭행렬 S 의 고유벡터 q 는 정규직교할 수 있다.
- ⑦ 모든 행렬의 특이벡터는 정규직교한다.

직교행렬 생성의 성공 여부는 벡터의 크기를 정의하는 제곱합($\|x\|^2 = x^T x$)과 깊이 연결되어 있다. 최소제곱 문제에서 $\|Ax - b\|^2$ 의 도함수는 대칭행렬 $S = A^T A$ 를 생성한다. S 는 직교행렬 Q 에 의해 대각화된다(6장에서 고유벡터에 대해 자세히 설명한다).

모든 행렬 A 는 두 직교행렬 Q_1, Q_2 에 의해 대각화된다(7장에서 특이벡터에 대해 자세히 설명한다). 직교행렬에 대해 추가로 설명한 내용은 $A = QS$ 와 $A = QR$ 이다.

- ⑧ 모든 가역행렬은 직교행렬 Q 와 대칭행렬 S 의 곱이다. 예를 들어 다음과 같다.

$$A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \sqrt{5} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = QS$$

- ⑨ 모든 가역행렬은 직교행렬 Q 와 상삼각행렬 R 의 곱이다. Q 와 R 은 4.4절의 그람-슈미트 과정으로부터 얻는다. 예를 들어 다음과 같다.

$$A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 0 & 3 \end{bmatrix} = QR$$