리드오프 2주차 후반부

행렬대수학 및 선형대수학에서 배운 기본적인 개념(벡터와 행렬의 연산, 곱, 행렬식, transpose 등)은 숙지하고 있다고 가정한 상태에서 교안을 작성했습니다. 저의 개인적인 바람은 앞으로 3번의 리드오프를 통해 학회원 분들이 선형대수학을 데이터적 관점에서 바라보고, 각 개념들이 통계학적으로는 어떤 의미로 확장될 수 있는지 파악하여 시야를 넓히는 것입니다. 그 과정에서 재미를 느낄 수 있다면 더욱 좋구요😊 만약 아래 개념들을 모른다면 다소 어려울 수 있으니, 아래의 개념이 생소하시다면 한번 훑어보고 오시는 것을 추천드립니다. 질문도 언제나 환영입니다!

[벡터 연산, 선형 독립과 종속, 기저, 대각 행렬, 대칭 행렬, 행렬 연산과 행렬 곱, 행렬식, 역행렬, transpose]

이번 주차에는 선형 변환과 함수의 개념으로서 선형대수학을 바라보는 관점을 다룹니다.

1. 선형대수학 기초

1-0. 선형

우선 다음과 같은 특성이 있는 함수 f를 선형이라고 한다. 당연히 행렬과 벡터의 기초 연산 및 선형 변환 모두 선형이기에 선형의 개념은 확실히 알아두고 넘어가자.

$$f(x) + f(y) = f(x + y)$$
$$cf(x) = f(cx)$$

1-1. 벡터 개념 다시 보기

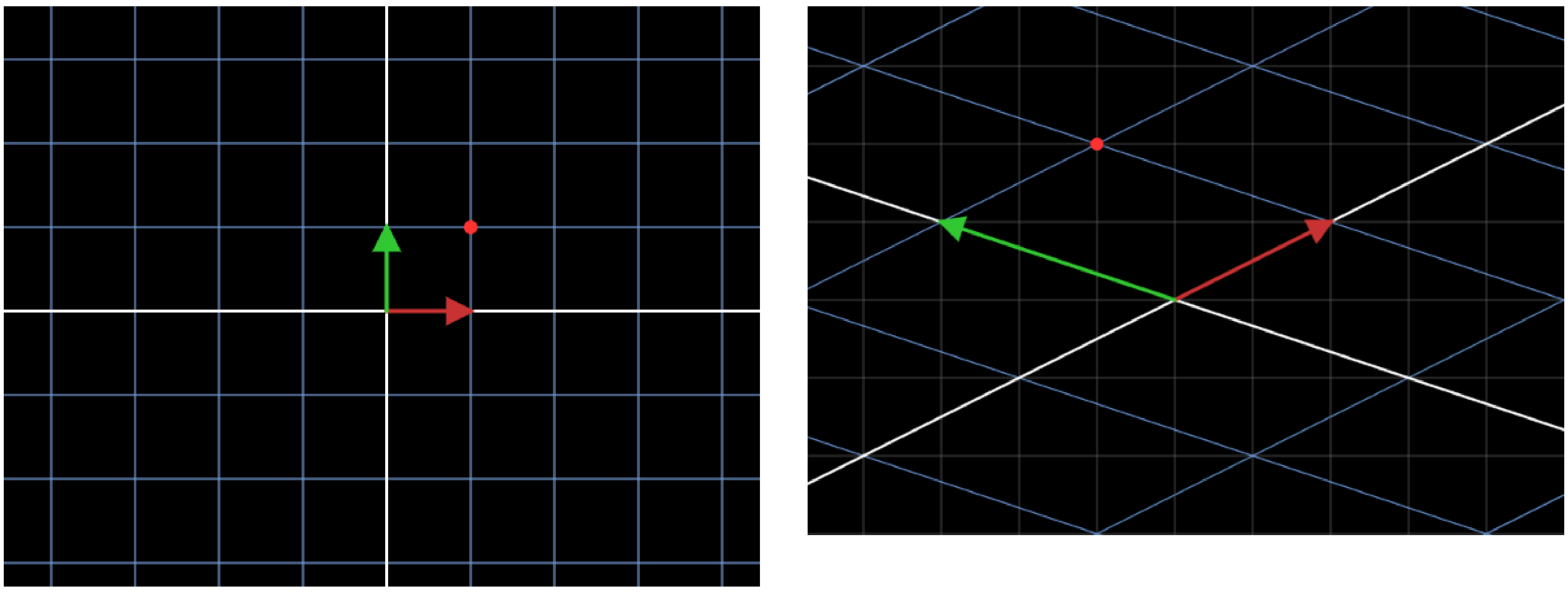
물리학에서 벡터란 '크기와 방향으로 정의되는 값'이라고 할 수 있다. 이것은 기하학적인 벡터의 특성을 잘 반영하고 있는 정의라고 할 수 있으며, 이전 시간 그래디언트를 위해 다루었던 벡터는 물리학적 특성을 띠고 있다고 할 수 있다.

그러나 우리는 통계학적 관점에서 선형대수학을 다루기 위해 벡터를 새로운 관점에서 바라볼 필요가 있다. 이는 그저 벡터를 벡터 공간(vector space)의 원소(점 하나), 또는 순서를 맞춰 숫자를 나열한 리스트정도로 바라보는 것이다. 이렇게 벡터를 바라보게 되면 벡터를 하나의 데이터 포인트로써 바라볼 수 있다. 전공에서 배우는 여러 가지 데이터 처리 기법들(PCA, SVD, 선형 회귀) 등의 수많은 기법들이 위와 같이 데이터를 벡터로 생각하며 데이터를 처리한다.

벡터의 각 성분을, "벡터를 어떻게 늘리고 줄일 것인지에 대한 정보"라고 보는 것도 다음에 나올 행렬과 선형 변환 개념을 이해하는데 도움이 됩니다. 예를 들어, $\begin{bmatrix}3\\2\end{bmatrix}$라는 벡터는 두 개의 표준 기저 벡터 $\begin{bmatrix}1\\0\end{bmatrix}$ $(\hat{i})$과 $\begin{bmatrix}0\\1\end{bmatrix}$ $(\hat{j})$을 각각 3과 2만큼 스케일링해 합한다고 보는 것입니다. 이러한 관점은 1-2를 이해하는데 도움이 될 수 있습니다.

1-2. 행렬 개념 다시 보기

이제 행렬을 살펴보자. 선형대수학의 이해를 위해 행렬을 바라보는 좋은 관점은 행렬을 **공간의 선형 변환**으로써 보는 것이다. 결론부터 말하자면, 벡터 공간의 선형 변환과 행렬은 본질적으로 같다. 행렬 연산을 먼저 기하학적으로 관찰한 뒤, 수식과 함께 이해해보자. $\begin{bmatrix}2 & -3\\1 & 1\end{bmatrix}$라는 행렬을 예로 들자면,



$(1, 1)$라는 빨간 점이 선형변환 이후 $(-1, 2)$로 이동한 것을 확인할 수 있다. 또 발견할 수 있는 점은, 표준 기저 벡터들($\hat{i}$, $\hat{j}$)도 선형 변환 후 이동했다는 것인데, 여전히 빨간 점은 그 변환된 $\hat{i}, \hat{j}$을 하나씩 더한 벡터로 표현된다는 것이다. $(-1, -2)$의 점도 어떻게 이동하는지 직접 살펴보자.

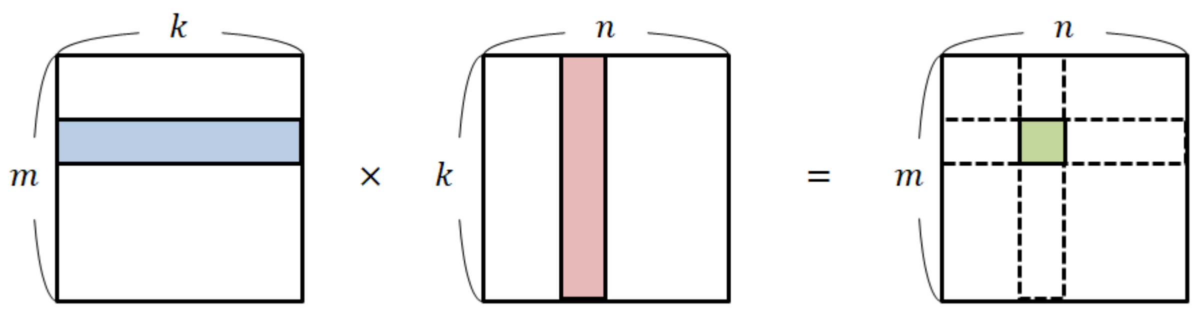
즉, 변환 전에 어떤 벡터를 이루는 $\hat{i}, \hat{j}$의 선형 결합이 변환 후에도 같은 선형결합을 유지한다는 것이다. 다시 표현하자면, $\hat{i}, \hat{j}$의 선형변환 후 위치만 알면, 알아보고자 했던 벡터의 위치도 추론이 가능하다는 것을 의미한다.

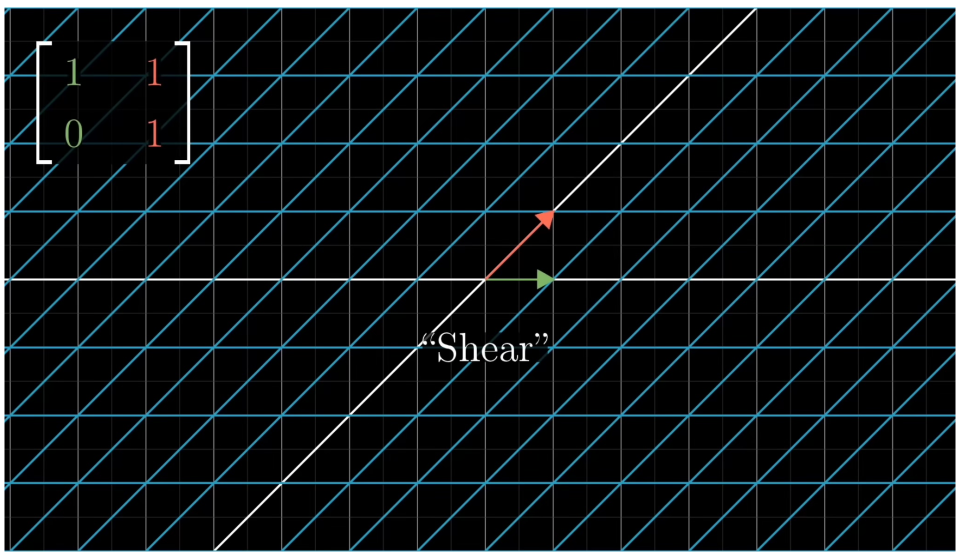
수식적으로는 $1\begin{bmatrix}1\\0\end{bmatrix}+1\begin{bmatrix}0\\1\end{bmatrix}$로 표현되었던 빨간 점은, 변환 후 $1\begin{bmatrix}2\\1\end{bmatrix}+1\begin{bmatrix}-3\\1\end{bmatrix}$의 점으로 바뀌었다. 이를 일반화하면 (x, y)라는 점은, $x\begin{bmatrix}2\\1\end{bmatrix}+y\begin{bmatrix}-3\\1\end{bmatrix}$이 된다.

$$\begin{bmatrix}\color{red}{2} & \color{blue}{-3}\\ \color{red}{1} & \color{blue}{1}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}2x-3y\\1x+1y\end{bmatrix}$$

이제 행렬의 연산으로는 다음과 같이 이해할 수 있다. 빨간색은 $\hat{i}$이 선형 변환 후 도착하는 장소이다. 파란색은 $\hat{j}$이 도착하는 장소이다. (x, y)는 원래 $\hat{i}$의 x배와 $\hat{j}$의 y배의 선형 결합으로 만들어지는 벡터였으므로, 선형 변환 후 (x, y)는 여전히 선형변환 후의 기저 벡터들이 같은 선형 결합으로 만드는 벡터가 된다.

즉, 변환 후 새 기저 벡터들로 스케일링하고 합한다는 개념으로 이해할 수 있다. 이것이 선형 변환이 의미하는 바이며, 조금은 신기하게 보였던 행렬의 곱이 이루어지는 원리이다.



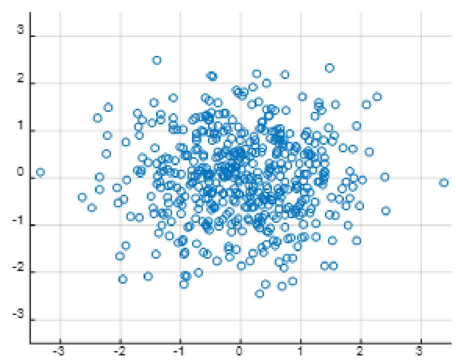


위 그림은 Shearing이라는 대표적 선형 변환이 기하학적으로 어떻게 보여질 수 있는지를 보여주는 예시이다. 회전, 대칭, 사영 등의 변환이 이루어지기 위해서는 기저 벡터가 어떻게 변환되어야할지 기하학적으로 상상한 후, 행렬대수학, 또는 선형대수학 교안을 통해 답을 확인하는 것도 정말정말 좋은 복습 및 이해의 방법이 될 것이니 꼭 해보도록 하자.

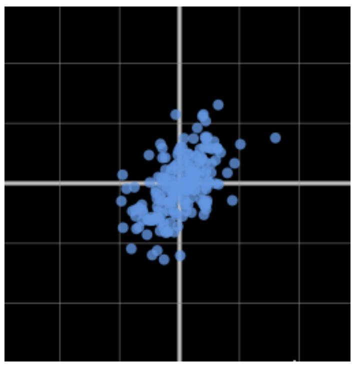
이 관점에서 행렬을 바라보게 되면 여러 가지 행렬의 기하학적 의미에 대해서도 이해가 가능해진다.

예제1) 공분산 행렬($\frac{X^TX}{N}$)

공분산 행렬은 데이터의 구조를 설명해주며, 특히 특징 쌍(feature pairs)들의 변동이 얼마나 닮았는가(얼마만큼이나 함께 변하는가)를 행렬에 나타내고 있다고 할 수 있다. 즉 조금 다르게 말하면, 우리는 지금 보고 있는 데이터의 분포에 대해 "원래 원의 형태(고른 형태)로 주어졌던 데이터가 공분산 행렬(선형변환)에 의해 변환된 결과"라는 관점에서 데이터를 보는 것이다.



다음과 같이 Gaussian distribution을 통해 랜덤하게 분포된 선형 변환 전 데이터의 분포를 생각해보자. 그리고 $\begin{bmatrix}3 & 2\\2 & 4\end{bmatrix}$라는 선형 변환으로 원래 데이터를 변환시키면 다음과 같다.

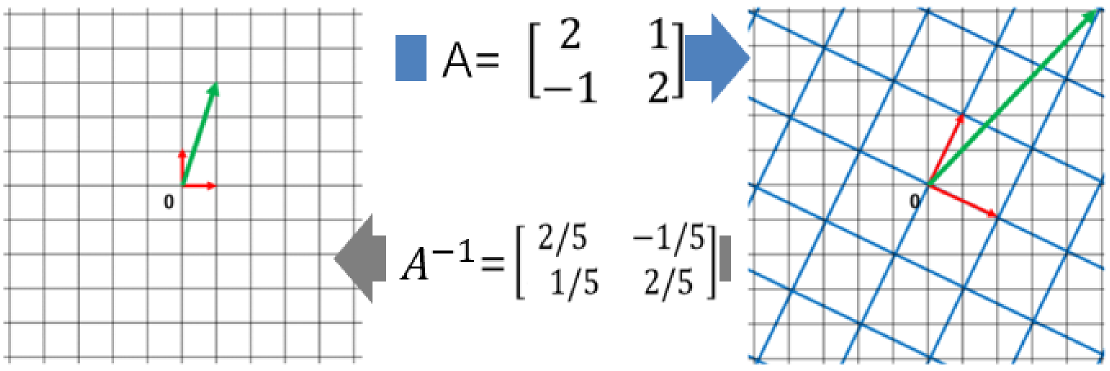


즉 선형 변환 $\begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix}$가 의미하는 것은 다음과 같다.

1. 1행 1열의 원소는 1번 feature의 분산을 나타낸다. 즉, x**축 방향으로 데이터가 얼마만큼 퍼지게 할 것인가**를 말해준다.
2. 1행 2열의 원소와 2행 1열의 원소는 각각 x, y축으로 **함께 얼마만큼 퍼지게 할 것인가**를 말해준다.
3. 2행 2열의 원소는 **y축 방향으로 데이터가 얼마만큼 퍼지게 할 것인가**를 말해준다.

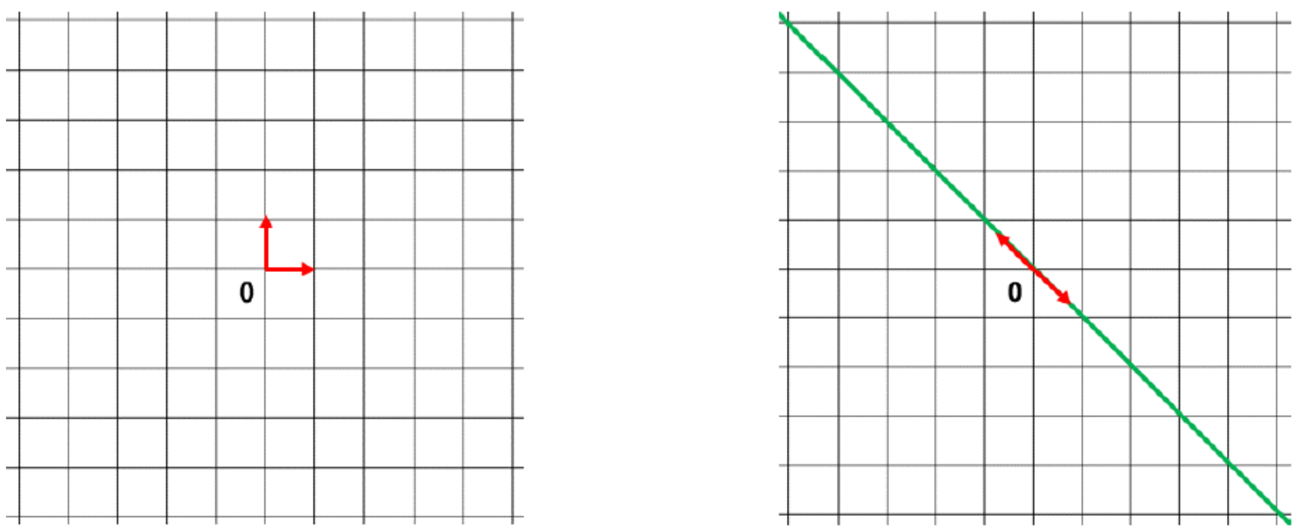
이제 우리는 공분산 행렬의 차원이 확장되어도 데이터가 어떻게 분포되었는지 짐작해볼 수 있을 것이다. $\begin{bmatrix} 3 & 2 & -5 \\ 2 & 4 & 3 \\ -5 & 3 & 8 \end{bmatrix}$ 의 공분산 행렬을 갖는 3차원 데이터에 대해서는 실제로 한번 떠올려보자. 선형대수학은 이렇게 우리가 표현할 수 없는 고차원에 대해 사고할 수 있도록 해준다는 점에서 의미가 있으므로, 차원을 늘려가며 생각해보는 과정도 중요하다.

예제2) 역행렬(A^{-1})



왼쪽의 좌표계에서 행렬 A를 곱하면 오른쪽의 좌표계로 선형 변환이 이뤄진다는 것은 전에서 살펴보았다. 그렇다면 이렇게 변환된 파란색 공간을 다시 원래의 공간으로 되돌리려면 파란색 좌표계를 기준으로 봤을 때, 공간을 왼쪽으로 회전하고 확장된 만큼 다시 축소시킨다면 원래의 좌표계로 돌아갈 수 있을 것이다. 이렇게 A의 linear transformation을 다시 되돌리는 행렬을 A의 역행렬이라고 한다. 즉 오른쪽의 좌표계를 원래의 좌표계로 보고, A의 역행렬을 곱해준다면? 다시 왼쪽의 좌표계로 돌리는 것이 가능해진다.

A의 역행렬이 존재한다는 것은, $Ax = b$ 가 유일한 해를 갖는다(unique) 와 equivalent하다는 것을 기억할 것이다. 즉 그 해가 유일하다는 것은 특정 x를 선형변환한 Ax가 유일하다는 것을 의미하며, 다시 일대일 대응으로 왼쪽의 벡터로 되돌릴 수 있음을 의미한다.

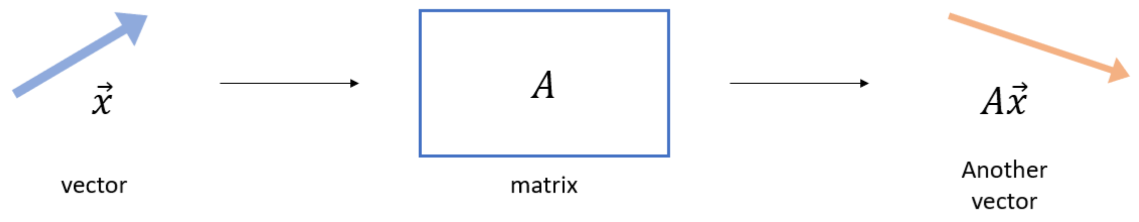


그러나 위 예시처럼 선형변환 전후의 차원이 달라진다면? 이 경우는 오른쪽의 계를 기본 계로 보았을 때, 그 1차원의 벡터를 어떤 특정 2차원 벡터로 되돌려야 하는지 알 수 없다. 즉, x와 Ax는 일대일 대응이 아니었음을 의미하며, 결국 $Ax = b$의 해가 유일하지 않다, 역행렬이 존재하지 않는다를 의미한다.

2. 함수와 4가지 주요 선형 공간

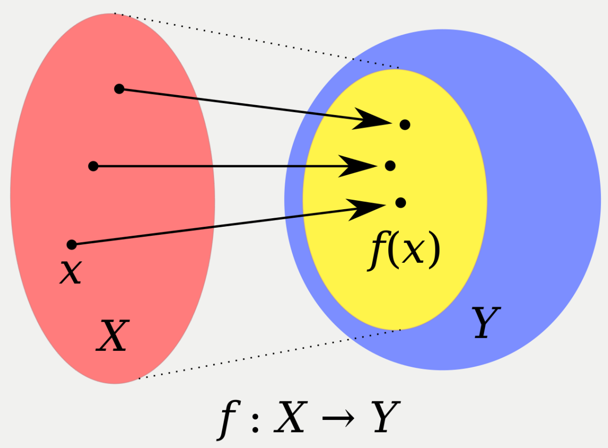
2-0. Introduction

이제 행렬이 선형 변환임을 알았으니, '행렬은 벡터를 변환시켜 다른 벡터를 출력해준다' 라는 관점에서 접근해볼 수 있다.



조금 다른 말로 하자면, 행렬이라는 것은 벡터를 입력 받아 벡터를 출력해주는 함수라고도 할 수 있을 것 같은데, 정말 선형 변환이 함수라면 함수의 근본적인 정의를 만족할 수 있는 것일까? 라는 점이 궁금해진다. 입력과 출력이 있다고 해서 다 함수라고 말할 수는 없는 것인데, 우리가 말하는 **'선형 변환'이 함수라고 엄밀하게 부를 수 있는 것일까** 를 알아보자.

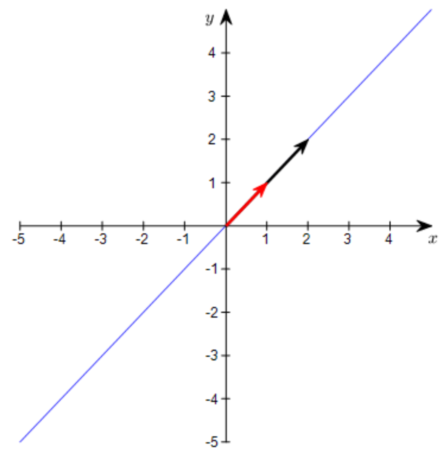
함수의 정의를 먼저 살펴보자.

> 📌 집합 X, Y에 대해 각각 정의역, 공역이라고 하자. 임의의 $x \in X$에 대하여, $(x, y) \in graph\ f$인 $y \in Y$가 유일하게 존재한다면, 이러한 y를 $f(y)$라고 한다.
>
> 즉 함수는 정의역의 각 원소를 정확히 하나의 공역 원소에 대응시킨다.
>
> 
>

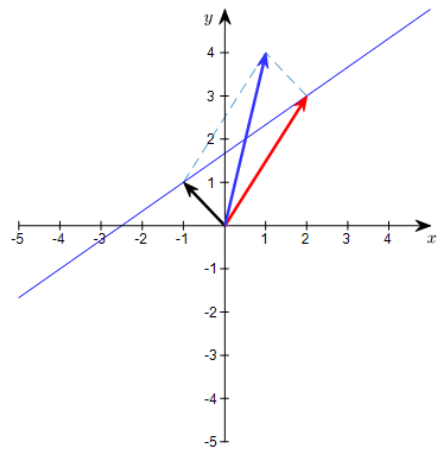
즉 우리가 선형변환을 함수로 생각할 수 있다고 한다면, 엄밀하게는 함수의 근본적인 의미인 정의역->치역의 매핑에 대해 생각해볼 수 있어야 한다. 그렇다면 선형변환에서 정의역, 공역, 치역은 각각 어떤 것일까? 그것이 바로 네 개의 주요 부분공간(row space, null space, column space, left null space)이 의미하는 바이다.

2-1. 부분 공간

부분 공간이라는 것은 부분 집합의 개념을 벡터 공간에 접목시킨 것으로 볼 수 있다. 즉 집합에서 부분 집합이 있는 것처럼 벡터 공간에서도 벡터 공간의 기본 구조를 그대로 유지하는 작은 벡터 공간, 즉 부분 공간이 있다. 가령, 2차원 실수 공간에서 부분 공간을 하나 생각해보자면 원점을 지나는 직선 상에 있는 모든 벡터들의 집합은 1차원 부분 공간을 이룬다.



원점을 지나는 직선
두 벡터의 덧셈 및 스칼라곱셈이
직선 내에 닫혀 있음
➔ 벡터 공간 $\mathbb{R}^2$의 부분 공간



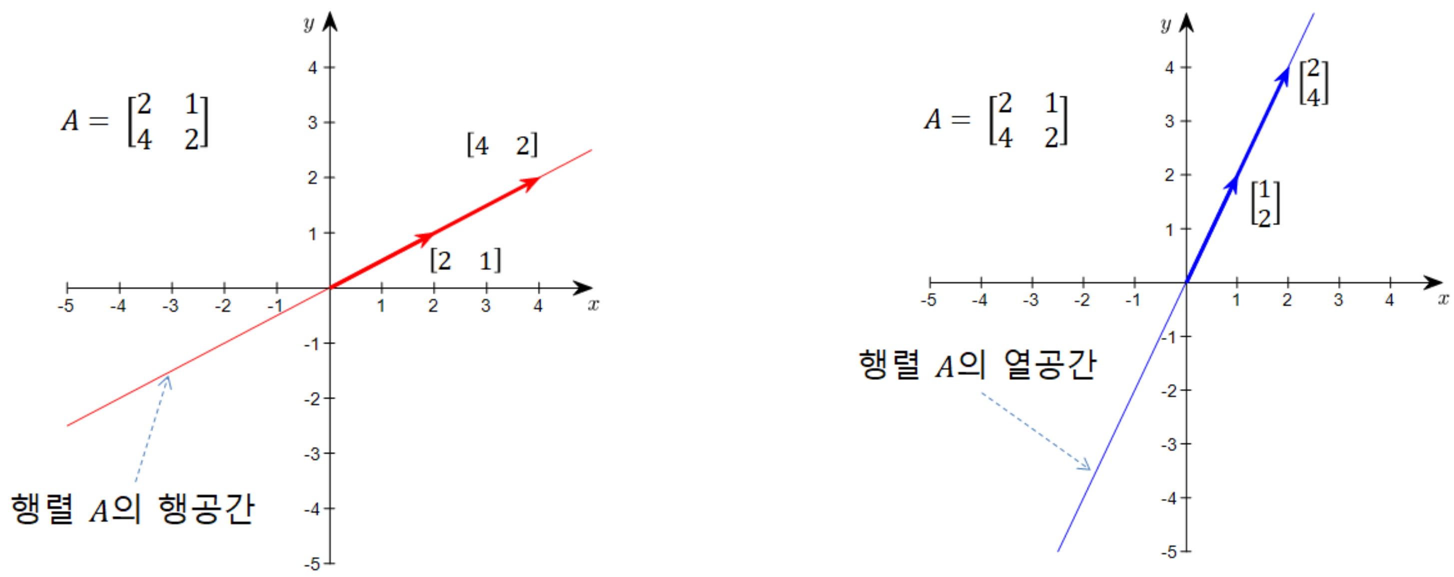
원점을 지나지 않는 직선
두 벡터의 덧셈 및 스칼라 곱셈이
직선 밖에 존재하게 됨
➔ 벡터 공간 $\mathbb{R}^2$의 부분 공간이 아님

1) 행공간과 열공간(Row Space, Column Space)

이런 관점에서 보았을 때, 우리에게 주어진 임의의 행렬 A의 모든 행 혹은 모든 열들의 선형 결합으로 구성된(span) 벡터공간은 부분 공간이며 각각을 행공간(row space), 열공간(column space)이라고 부른다. 예를 들어 행렬 A가 아래와 같이 주어져있다고 해보자.

$$A = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}$$

그러면, 행공간은 벡터 $\begin{bmatrix} 2 & 1 \end{bmatrix}$, $\begin{bmatrix} 4 & 2 \end{bmatrix}$ 의 선형 결합으로 이루어진 선 상에 있는 모든 벡터들의 집합이다. 열공간은 반대로 $\begin{bmatrix} 2 & 4 \end{bmatrix}^T$와 $\begin{bmatrix} 1 & 2 \end{bmatrix}^T$의 선형 결합으로 이루어진 선 상에 있는 모든 벡터들의 집합일 것이다. 열공간은 이전 챕터에서 보았던 변환된 기저벡터가 만드는 공간 으로 이해해보는 것도 좋다.



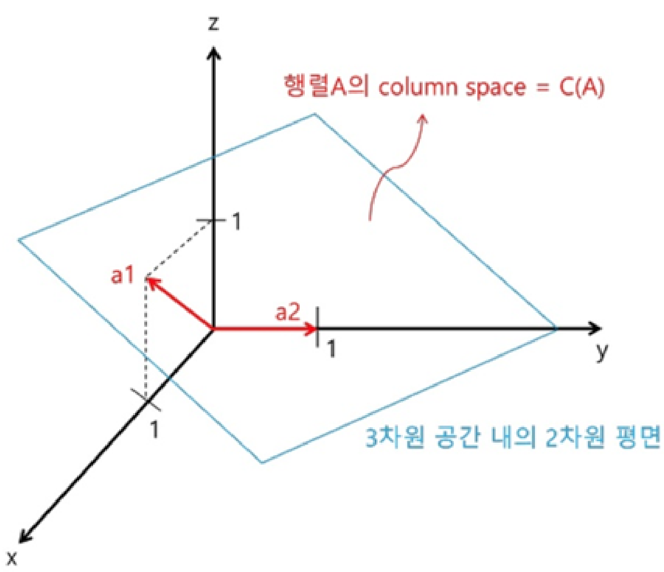
위 예시에서 A행렬의 두 벡터가 선형 종속이었기 때문에, 열공간과 행공간이 직선으 로 표현되었다. 그렇다면 3×2 행렬로 확장해보자.

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

이 행렬로 $Ax = b$ 문제를 푼다고 가정하면, 다음과 같이 행렬 A의 열벡터의 선형결합으로 식이 만들어진다.

$$x_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

위 선형 결합으로 만들 수 있는 공간에서 b 를 구할 수 있게 되고 따라서 벡터 b 의 첫 번째 성분과 세 번째 성분이 같을 때만 x_1, x_2를 구할 수 있게 된다. 이는 3차원 공간에서 어떤 2차원 평면을 만들게 될 것이다.

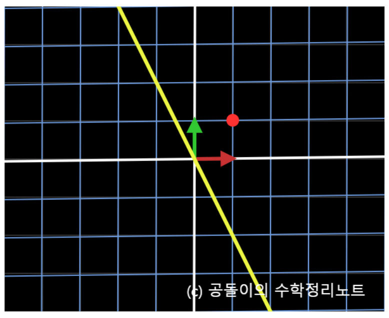


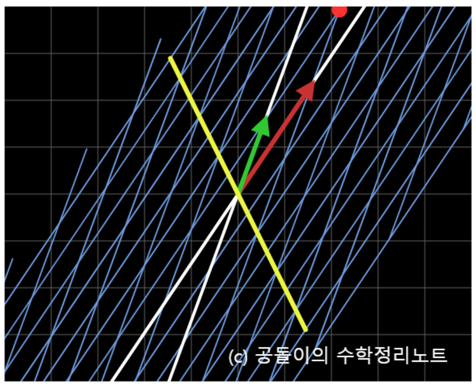
이를 행렬 $A_{m \times n}(m > n)$로 일반화하면, 선형독립인 열벡터의 개수가 $r(\leq n)$일 때 열공간은 R^m 차원 내부에서 r차원 초평면을 구성한다. 열공간의 차원(dimension), 즉 변환 결과의 차원은, 추후 Rank라는 개념과 연관되게 된다.

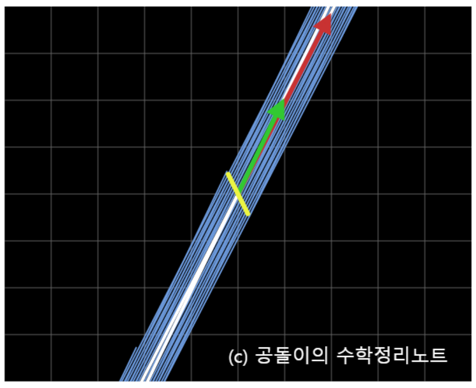
2) 영공간(Null Space)과 Left Null Space

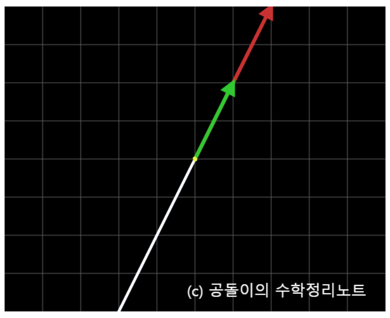
행렬 A에서 즉각적으로 인지하긴 어렵지만 영공간(null space)라는 부분 공간도 존재한다. 영공간은 $Ax = 0$ 의 조건을 만족하는 $\vec{x}$ 들의 집합이다. 즉, A라는 선형 변환 후에 모두 0을 출력하게 만들어주는 입력 벡터 x 들인 것이다. 행렬 $A = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}$일 때, A라는 선형변환이 어떻게 작동하는지 다음의 그림을

통해 보자.









위 과정을 보면, 2차원 벡터 공간상에 있던 모든 벡터가 열공간으로 이동하는 것을 확인할 수 있다. 이 중 선형 변환 후에 $(0,0)$으로 이동하는 점들의 집합이 바로 노란색 선이고 이것이 바로 **영공간(Null Space)**이다. 이때, 영공간과 행공간은 서로 직교하는 공간이다.

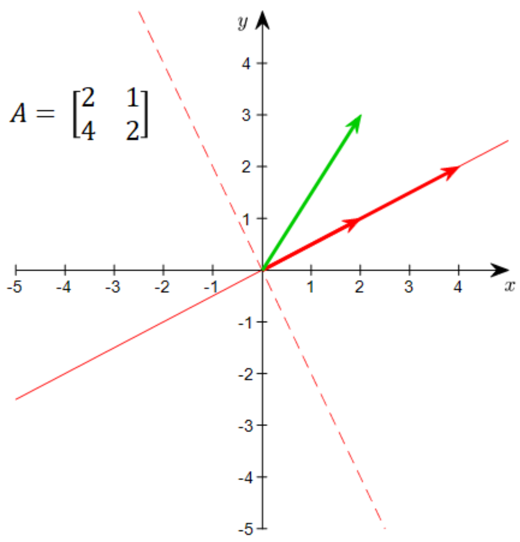
Left Null Space는 행렬 A^T의 영공간이다. Left Null Space는 선형 변환 과정에서 시각화할 수는 없지만, 열공간과 서로 직교한다는 점 정도만 기억해두자.

2-3. Fundamental Theorem of Linear Algebra

그러면 이제 우리는 비로소 제대로 선형 변환을 함수로써 바라볼 수 있게 된다. 정리하자면 정의역과 공역의 집합을 벡터 공간으로 봤을 때, A가 $m \times n$ 행렬이라면 n차원 벡터 공간이 정의역이 되고 m차원 벡터 공간이 공역이 되는 것이다.

1) 입력(정의역): row space + null space = $\mathbb{R}^n$

선형 변환의 정의역은 row space와 null space의 합집합이다. n차원 실수 공간 상의 어떤 벡터라도 row space와 null space 상의 벡터들의 선형조합으로 표현할 수 있다. 2×2 차원의 행렬, $A = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}$에 대해 생각해보자. $[2,3]$이라는 벡터는 행공간 위의 점(벡터)도 아니고, 영공간 위의 점(벡터)도 아니다. 다만, 행공간과 영공간이 서로 직교한다는 사실을 이용해 $[2,3]$이라는 벡터를 행공간과 영공간의 기저들의 선형결합으로서 표현할 수 있게 된다.



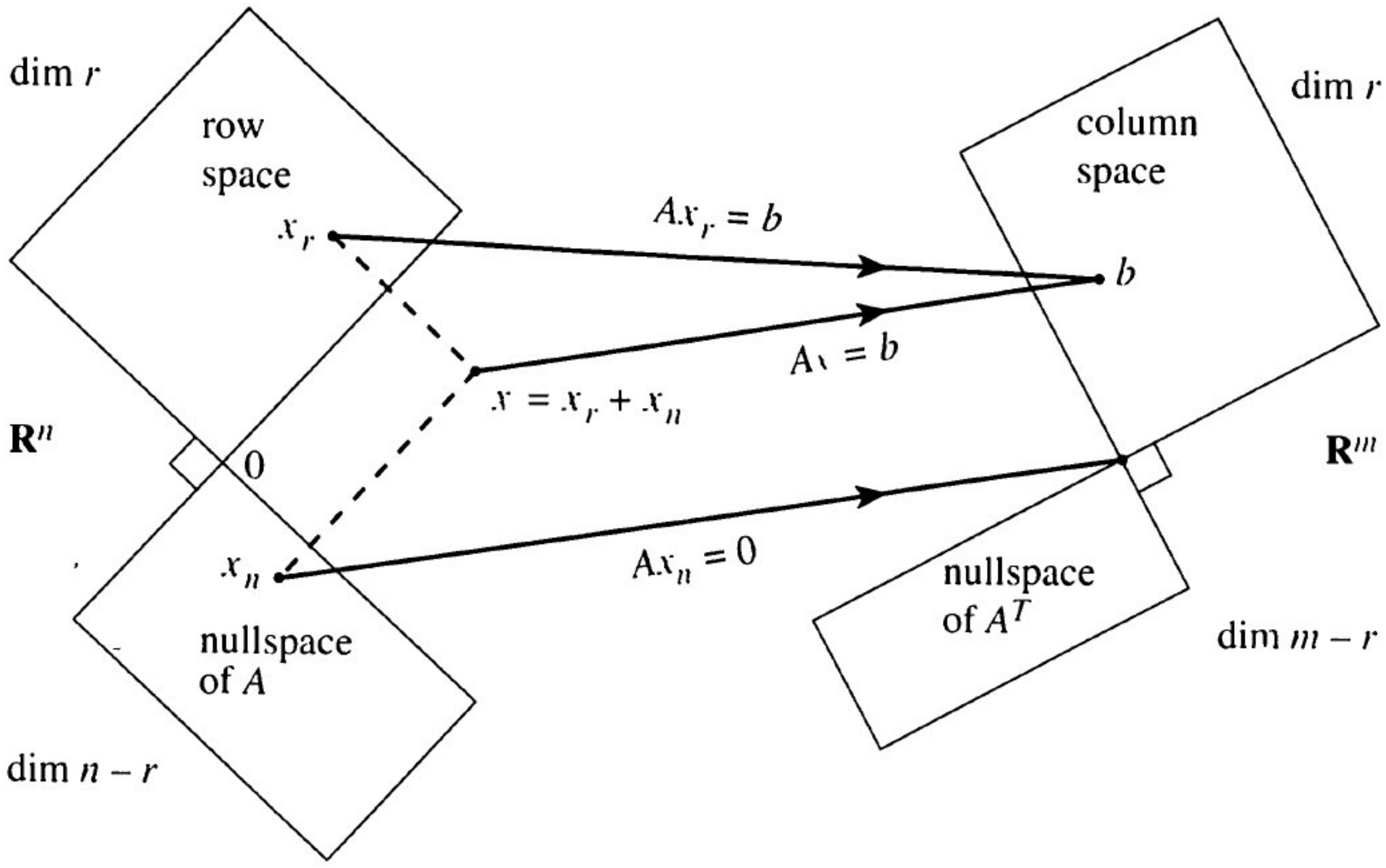
2) 출력(치역): 모든 것이 column space로

모든 정의역에 있는 벡터들은 열공간 위의 점으로 변환되게 된다. 그 이유는 1-2에서 보았던 것처럼 행렬과 벡터의 곱은 열벡터의 선형결합으로 표현될 수 있기 때문이다. 또 정의역의 벡터들은 모두 행공간의 기저와 영공간의 기저의 선형결합으로 구성되는데, 영공간의 기저로 표현되었던 벡터의 원소들은 모두 선형 변환 후 그 크기가 0으로 줄어들기 때문에 선형 변환 후에 모든 벡터들이 열공간 위에 위치하게 되는 것이다.

3) 공역: m 차원 실수 공간

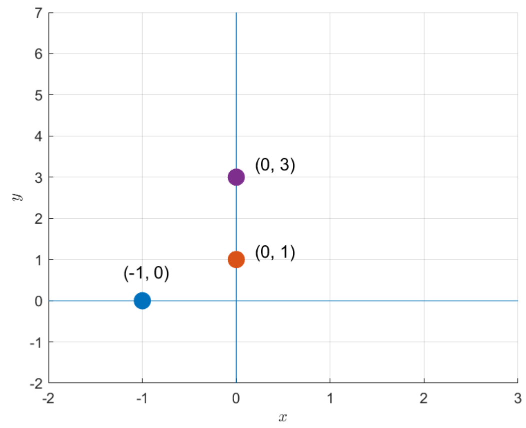
선형 변환의 치역은 column space이다. 공역에서 치역을 뺀 것이 left null space이다. 선형변환이라는 함수에서 공역은 column space + left null space 이며, column space와 left nullspace는 서로 직교한다.

이 주요 부분 공간들과 그 관계를 정리하면 아래, 선형대수학을 공부하며 한번쯤은 보았을 유명한 그림으로 표현할 수 있다.



3. 회귀분석의 Concept과 선형대수학

다음과 같이 세 개의 데이터 포인트가 주어져 있다고 하자.



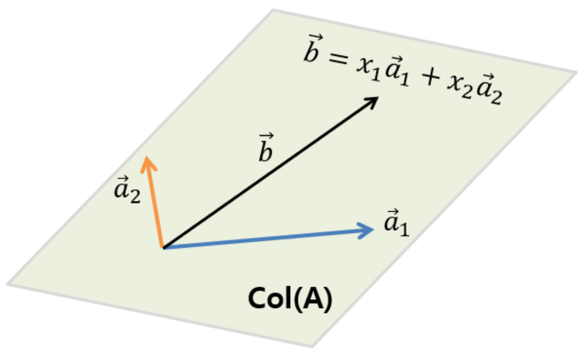
당연하게도, 2차원 평면 상에 어떻게 선을 놓더라도 이 세 점을 동시에 통과하는 직선을 구할 수는 없다. 즉 이 문제는 풀릴 수 없다. 해가 존재하지 않기 때문이다. 이는 미지수의 개수보다 데이터가 많기 때문이라고도 할 수 있고, 또 $Y = X\beta$의 해를 구할 수 없기 때문이라고도 할 수 있다.

$$(Y = X\beta) \Rightarrow \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_0 \end{bmatrix}$$

여기서 벡터와 행렬을 모두 열벡터로 표현하면 아래와 같이 표현되는데,

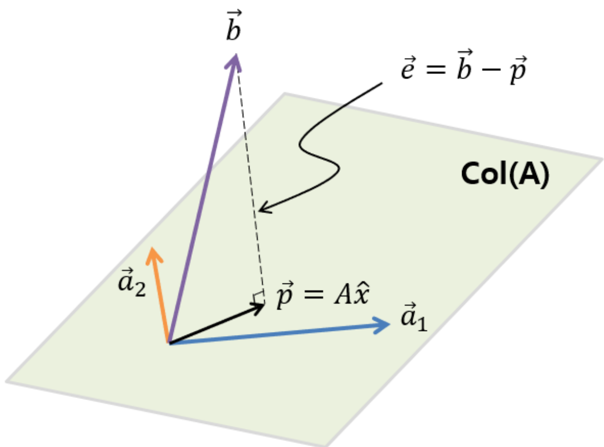
$$\Rightarrow \begin{bmatrix} | & | \\ \vec{a}_1 & \vec{a}_2 \\ | & | \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_0 \end{bmatrix} = \begin{bmatrix} | \\ \vec{y} \\ | \end{bmatrix} \Rightarrow \beta_1 \begin{bmatrix} | \\ \vec{a}_1 \\ | \end{bmatrix} + \beta_0 \begin{bmatrix} | \\ \vec{a}_2 \\ | \end{bmatrix} = \begin{bmatrix} | \\ \vec{y} \\ | \end{bmatrix}$$

즉, 열벡터 $\vec{a}_1$ 과 $\vec{a}_2$를 어떻게 조합하면 $\vec{y}$ 를 얻어낼 것인가? 라는 물음에 적절한 조합 비율인 $\beta_1(x1)$과 $\beta_0(x2)$를 찾는 것과 같은 이야기인 것이다. 따라서 해가 존재하기 위해 $\vec{a}_1$ 과 $\vec{a}_2$의 생성 공간(**열공간 = 치역**) 안에 $\vec{y}$ 가 포함되어 있어야 하는데, 이 경우는 그러한 경우가 아니다.



일반적으로 많은 데이터에 대해 그런 완벽한 직선은 존재하지 않는다. 그렇다면 최대한 가까운 직선이 그나마 괜찮은 답이 될 수 있지 않을까? 바로 이것이 최소제곱법의 아이디어이다. 이는 선형대수학의 관점에서 생각했을 경우 열공간안에 있는 정답에 가장 가까운 해를 찾는 과정과 일치시켜 생각할 수 있다.

그리고 아래의 그림에서 확인할 수 있듯이, 여기서 우리가 찾을 수 있는 $\vec{y}$ 와 가장 가까우면서, $\vec{a_1}$ 과 $\vec{a_2}$ 의 선형결합을 통해 얻을 수 있는 최적의 벡터는 $\vec{y}$ 가 열공간($col(A)$)에 정사영된 $\vec{p}$ 이다. 이제 우리는 이 $\vec{p}$ 를 통해 벡터 $\vec{a_1}$과 $\vec{a_2}$ 를 얼마만큼 선형 조합해야할지 알 수 있게 된다.



정리하자면, 원래의 목표는 $Y = X\beta$를 만족하는 β를 찾는 것이다. 해가 존재하기 위해서는 y가 X의 열공간에 있어야 한다. 하지만 대부분의 경우 그렇지 않다. 우리는 y를 X의 열공간에 가장 가깝게 근사시켜야 하고 이때 투영 행렬 H가 사용된다. H는 y를 변환하여 X의 열공간에 투영시키는 함수이다. 즉 이 과정을 통해 $HY = \mathbf{X}\hat{\beta}$ 를 만족시키는 해 $\hat{\beta}$의 해를 구할 수 있도록 하는 것이다. (보다 자세한 내용은 회귀팀 클린업 1주차 예정)

주요 선형 부분공간과 회귀분석

위 그림에서 $\vec{e}$는 열공간 상에 있는 모든 벡터들과 직교한다. 이를 행렬의 주요 부분공간 내용과 연결지어 생각하면 $\vec{e}$는 left null space에 있는 벡터임을 알 수 있다. 즉, $\vec{b}$는 열공간 상에서 만들 수 있는 기저벡터들과 left null space 상에서 만들 수 있는 기저벡터들을 합친 기저벡터들로만 구성할 수 있는 공간에 위치하고 있으며 $\vec{p}$는 그 중 가장 가까운 열공간 상에 있는 벡터, $\vec{e}$는 left null space 상에 있는 벡터를 의미하게 된다. 그림으로 표현하면 아래와 같다.

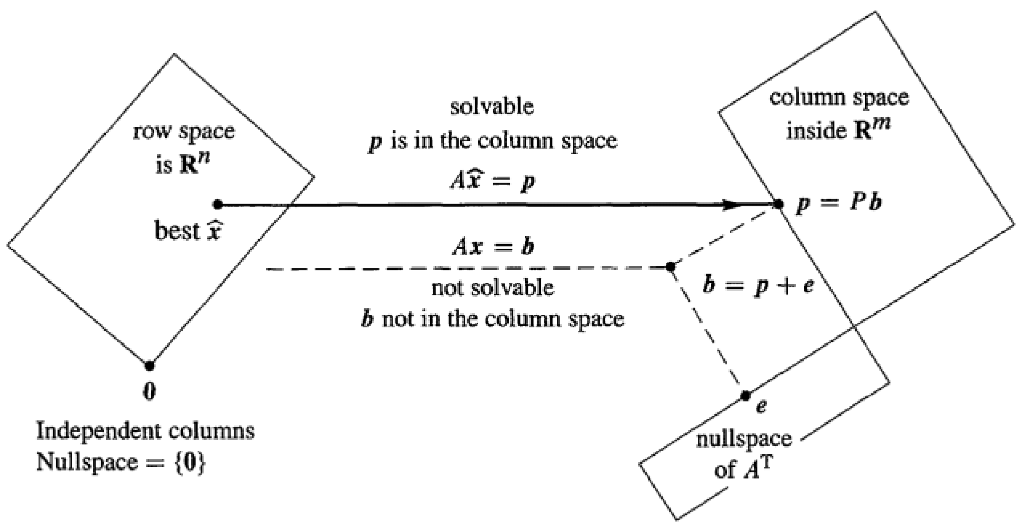
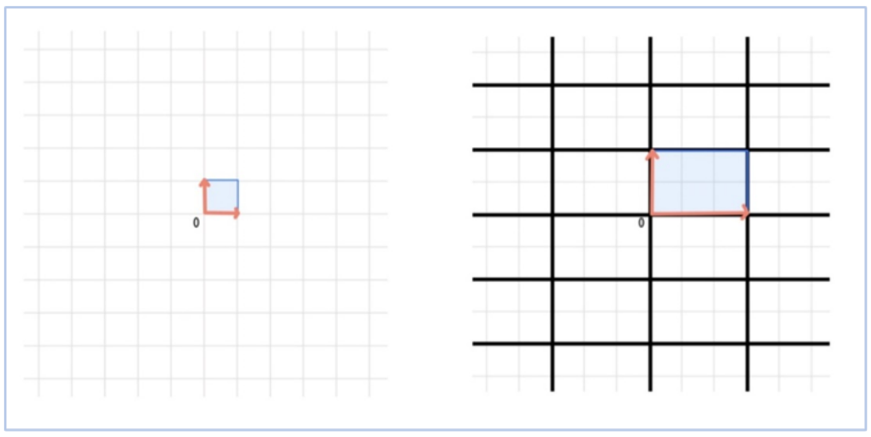


Figure 4.7: The projection $\boldsymbol{p} = A\hat{\boldsymbol{x}}$ is closest to $\boldsymbol{b}$, so $\hat{\boldsymbol{x}}$ minimizes $E = \|\boldsymbol{b} - A\boldsymbol{x}\|^2$.

Additional Topic - Determinant

행렬식은 정사각행렬에 수를 대응시키는 것으로, 이는 선형변환 후 공간이 얼마나 확장되거나 축소되는지에 대한 정보를 알려준다.



위 예시의 왼쪽 그림의 기저 벡터 $\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}0\\1\end{bmatrix}$ 를 두 변으로 해서 만들어지는 평행사변형의 넓이는 1이다. 오른쪽과 같이 $\begin{bmatrix}1\\0\end{bmatrix}$ 벡터 방향으로 2배, $\begin{bmatrix}0\\1\end{bmatrix}$벡터 방향으로 3배 확장시키는 선형변환을 가했을 때, 이 선형변환을 나타내는 행렬은 $\begin{bmatrix}2 & 0\\0 & 3\end{bmatrix}$ 일 것이다. (왜 그런지 이해가 간다면, 이제 행렬과 선형 변환을 기하학적으로 잘 이해했다고 볼 수 있다) 그리고 그 변형된 공간에서 두 기저 벡터를 변으로 하여 만들어지는 평행사변형의 넓이는 6일 것이다. 즉, 공간의 크기가 6배로 확장된 것을 알 수 있다. 이처럼 Determinant는 R^2 공간에서는 넓이와, R^3 공간에서는 부피와 상관이 있게 된다.

그렇게 되면

- $det\ AB = det\ A \times det\ B$

B에서의 변화율에 A에 의한 변화율이 차례대로 적용된다고 생각했을 때 위의 식이 성립된다고 이해할 수 있다.

- $det\ A = 0$

앞에서 determinant 가 R^2 공간에서는 넓이, R^3 공간에서는 부피와 관련이 있다고 했는데, 넓이나 부피가 0이 되는 행렬은 **공간을 압축시키는** 행렬, 즉 앞에서 보았던 역행렬이 없는 행렬의 경우라고 볼 수 있을 것이다. 이를 통해서 'A의 역행렬이 존재하지 않는다'와 'A의 determinant의 0'이라는 말 또한 equivalent하다는 걸 알 수 있다.