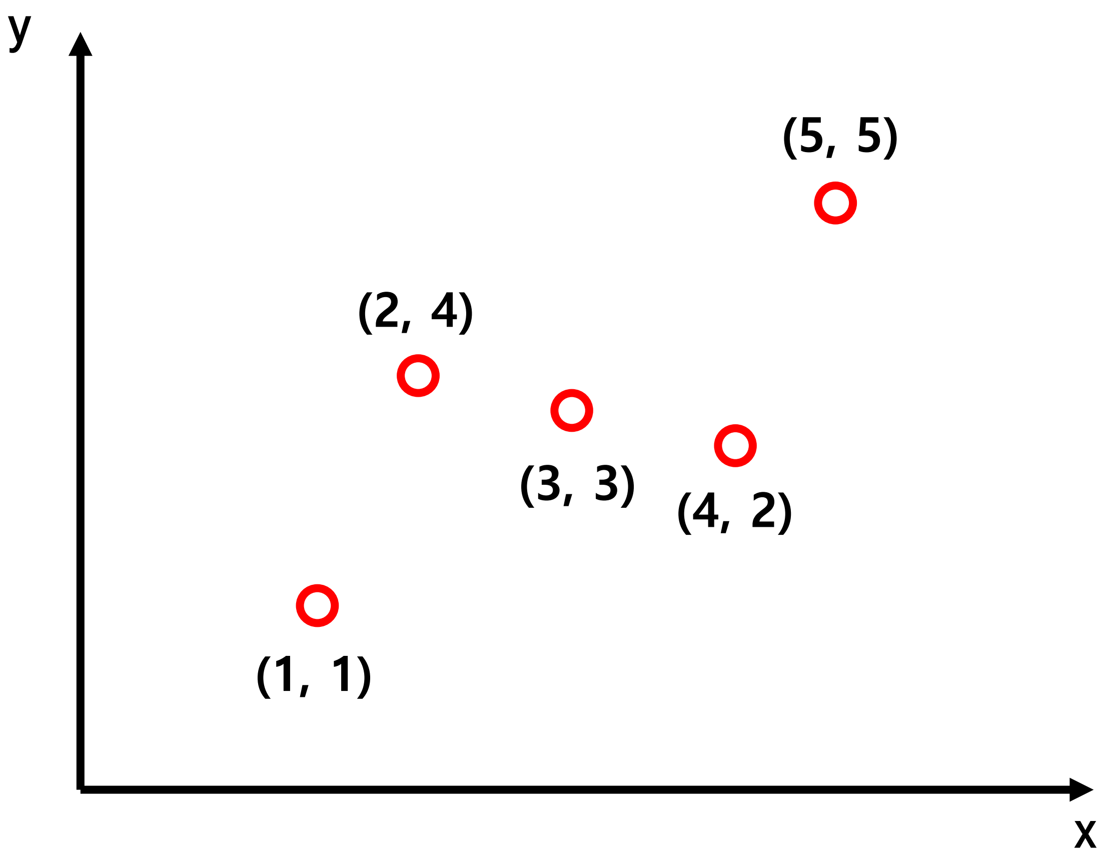
Least Squares

Prof. Jaewook Lee
Statistical Learning and Computational Finance Lab.
Department of Industrial Engineering
jaewook@snu.ac.kr
http://slcf.snu.ac.kr

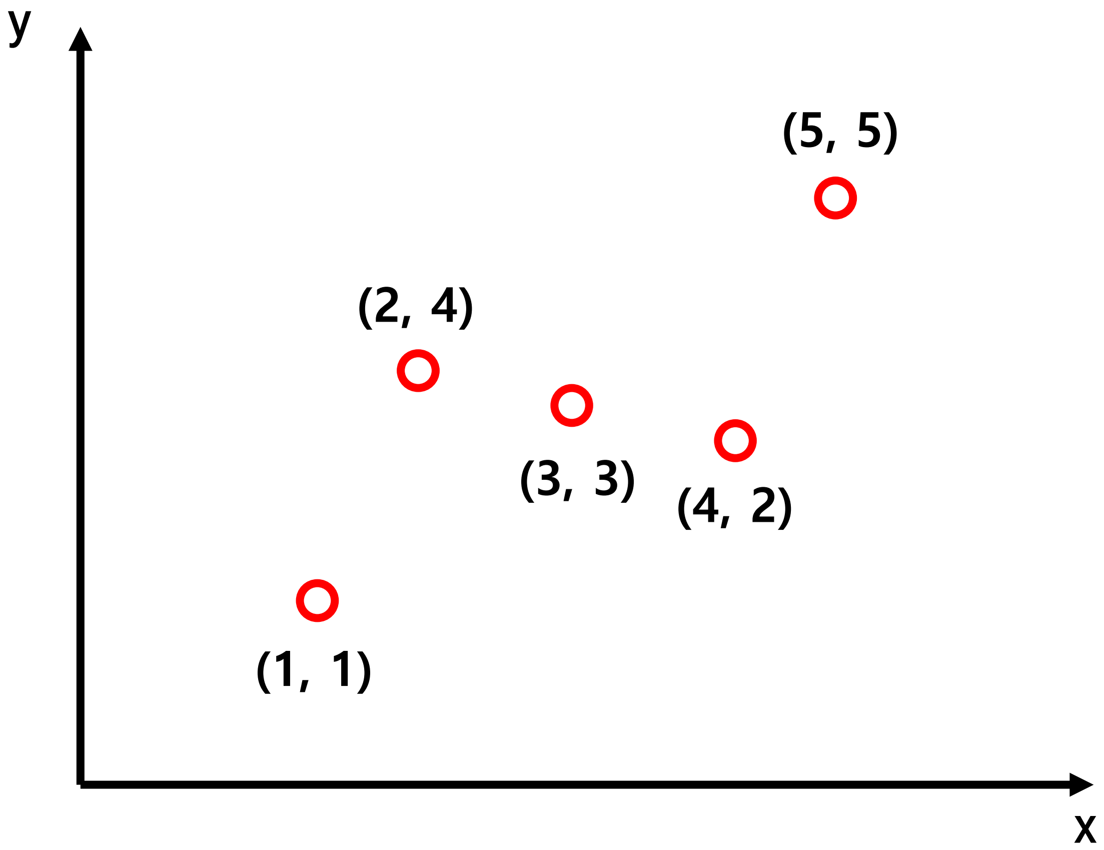
The Method of Least Squares

- 다음과 같은 데이터들을 가정해보자



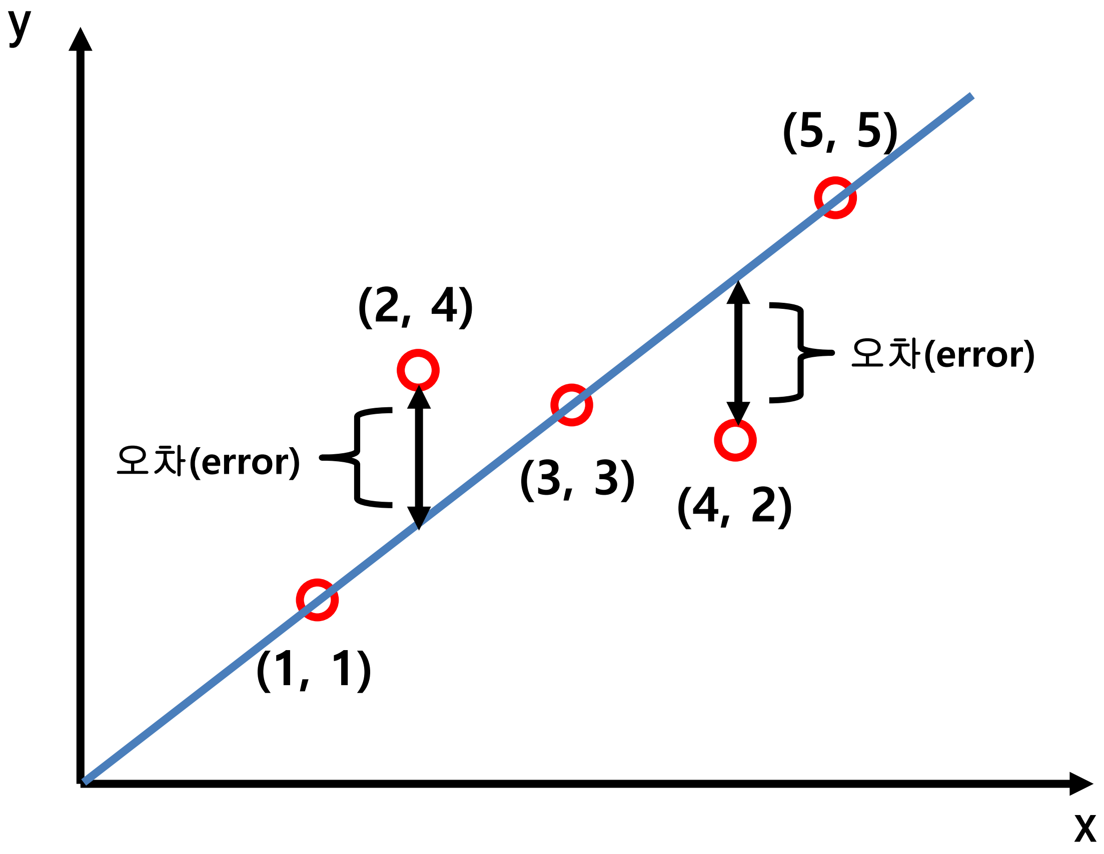
The Method of Least Squares

- 다음과 같은 데이터들을 가정해보자
  - 어떻게 하면 다음을 잘 설명하는 선형 모델을 만들 수 있는가?



The Method of Least Squares

- 다음과 같은 데이터들을 가정해보자
  - 어떻게 하면 다음을 잘 설명하는 선형 모델을 만들 수 있는가?



The Method of Least Squares

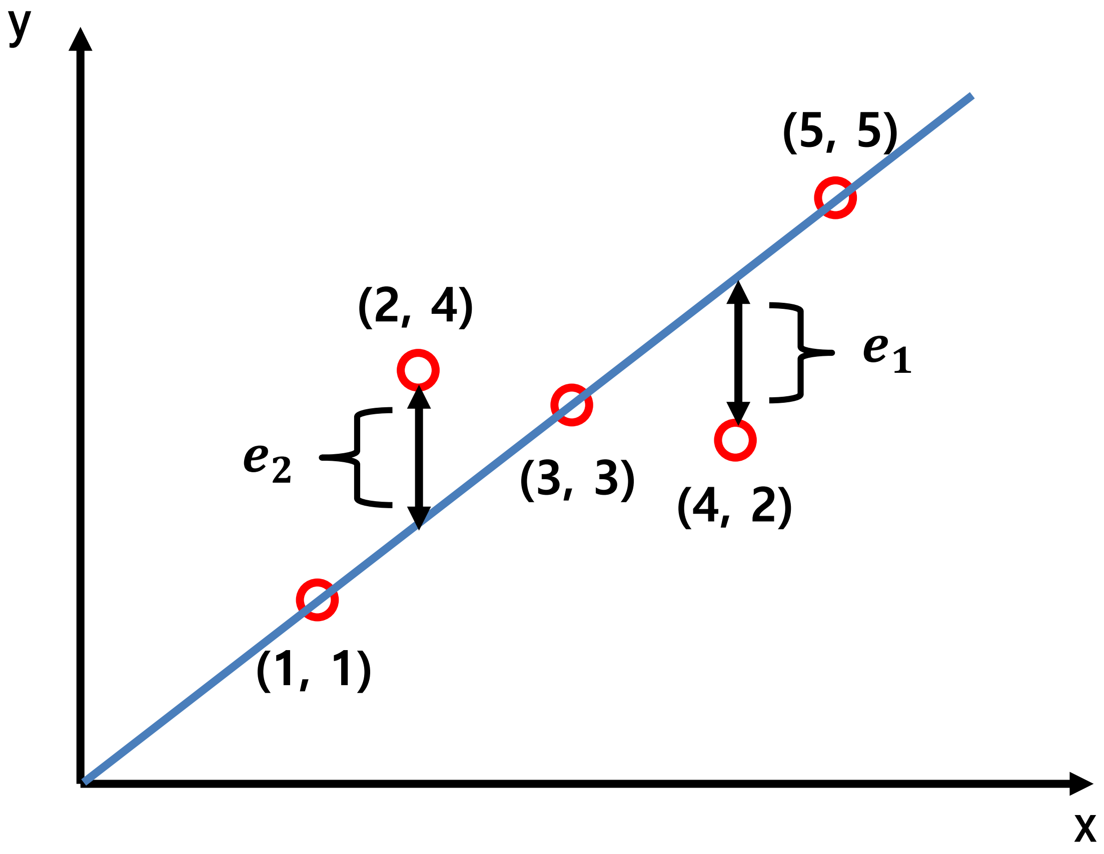
- Theorem (Least Squares Solution)
  - 여태까지의 선형 식은 아래와 같은 식으로 나타낼 수 있음

$$\begin{bmatrix}1\\2\\3\\4\\5\end{bmatrix}x=\begin{bmatrix}1\\4\\3\\2\\5\end{bmatrix}$$

  - 허나 위 식을 만족하는 x는 없음
  - 그렇다면?

The Method of Least Squares

- 다음과 같은 데이터들을 가정해보자
  - 좋은 모델임을 어떻게 말할 수 있는가?



The Method of Least Squares

- Theorem (Least Squares Solution)
  - 여태까지의 선형 식은 아래와 같은 식으로 나타낼 수 있음

$$Ax - b = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} x - \begin{bmatrix} 1 \\ 4 \\ 3 \\ 2 \\ 5 \end{bmatrix}$$

  - $min||Ax - b||$ 를 최소화하자!

The Method of Least Squares

- Theorem (Least Squares Solution)
  - $min||Ax - b||$
  - 목적함수를 최소화하는 것의 의미?

Four Fundamental Subspaces

- 정의

  - $A : \ m \times n \, matrix$
  - $\boldsymbol{R(A)}$ **: column space of** $\boldsymbol{A}$
  - $\boldsymbol{N(A)}$ **: null space of** $\boldsymbol{A}$
  - $\boldsymbol{R(A^T)}$ **: row space of** $\boldsymbol{A}$
  - $\boldsymbol{N(A^T)}$ **: left null space of** $\boldsymbol{A}$

Four Fundamental Subspaces

- 정의
  - A : $m \times n\ matrix$
  - $\boldsymbol{R}(\boldsymbol{A})$ **: column space of $\boldsymbol{A}$** = All linear combinations of the columns of A
  - $\boldsymbol{N}(\boldsymbol{A})$ **: null space of $\boldsymbol{A}$** = All linear combinations of x such that $Ax = 0$
  - $\boldsymbol{R}(\boldsymbol{A}^{\boldsymbol{T}})$ **: row space of $\boldsymbol{A}$** = All linear combinations of the rows of A
  - $\boldsymbol{N}(\boldsymbol{A}^{\boldsymbol{T}})$ **: left null space of $\boldsymbol{A}$** = All linear combinations of y such that $A^T y = 0$

Four Fundamental Subspaces

- 정의
  - $A : \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$
  - $\boldsymbol{R}(\boldsymbol{A})$: **column space of $\boldsymbol{A}$** $= a\begin{bmatrix} 1 \\ 3 \end{bmatrix} + b\begin{bmatrix} 2 \\ 4 \end{bmatrix}$
  - $\boldsymbol{N}(\boldsymbol{A})$: **null space of $\boldsymbol{A}$** $= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = 0 \rightarrow a + 2b = 0 \;\&\; 3a + 4b = 0$
  - $\boldsymbol{R}(\boldsymbol{A}^{\boldsymbol{T}})$: **row space of $\boldsymbol{A}$** $= a\begin{bmatrix} 1 \\ 2 \end{bmatrix} + b\begin{bmatrix} 3 \\ 4 \end{bmatrix}$
  - $\boldsymbol{N}(\boldsymbol{A}^{\boldsymbol{T}})$: **left null space of $\boldsymbol{A}$** $= \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = 0 \rightarrow a + 3b = 0 \;\&\; 2a + 4b = 0$

Four Fundamental Subspaces

- 정의
  - A : $\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$
  - $\boldsymbol{R}(\boldsymbol{A})$: **column space of** $\boldsymbol{A}$ $= a\begin{bmatrix} 1 \\ 3 \end{bmatrix} + b\begin{bmatrix} 2 \\ 6 \end{bmatrix}$
  - $\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 6 \end{bmatrix}$은 상수 배가 아닌가?

Linear Dependency

- 선형 종속(Linear dependent)
  - $v_1, v_2, \dots, v_k$에 대해 $c_1 v_1 + c_2 v_2 + \cdots + c_k v_k = 0$이 $c_1 = c_2 = \cdots = c_k = 0$일때만 존재한다면 위 벡터들은 서로 **선형 독립(linearly independent)**이라고 한다. 그렇지 않으면 **선형 종속 (linearly dependent)**라고 한다
  - $\begin{bmatrix}1\\3\end{bmatrix}, \begin{bmatrix}2\\6\end{bmatrix}$의 경우 $2 * \begin{bmatrix}1\\3\end{bmatrix} - 1 * \begin{bmatrix}2\\6\end{bmatrix} = 0$이므로 선형 종속이다
  - $span(v_1, v_2, \dots, v_k)$: $c_1 v_1 + c_2 v_2 + \cdots + c_k v_k$로 표현 가능한 모든 벡터이다

Linear Dependency

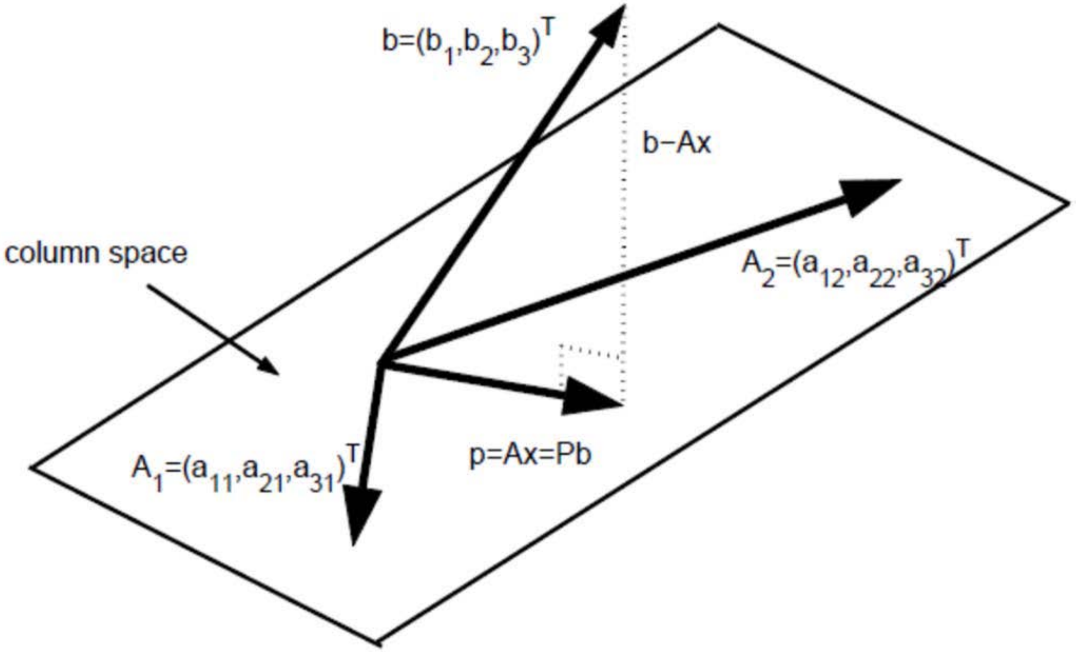
- 선형 종속(Linear dependent)
  - $A : \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$
  - $\boldsymbol{R}(\boldsymbol{A})$ **: column space of** $\boldsymbol{A} = a\begin{bmatrix} 1 \\ 3 \end{bmatrix} + b\begin{bmatrix} 2 \\ 6 \end{bmatrix} = a\begin{bmatrix} 1 \\ 3 \end{bmatrix} = span\{\begin{bmatrix} 1 \\ 3 \end{bmatrix}\}$

Least Squares

- $Ax = b$ 의 의미
  - $A : \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$
  - $Ax = b$ 의 의미는 무엇인가?
  - $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} x_1 + \begin{bmatrix} 2 \\ 4 \end{bmatrix} x_2 = b$
  - 즉, $\boldsymbol{b} \in \boldsymbol{R}(\boldsymbol{A})$이여야 해가 존재한다.

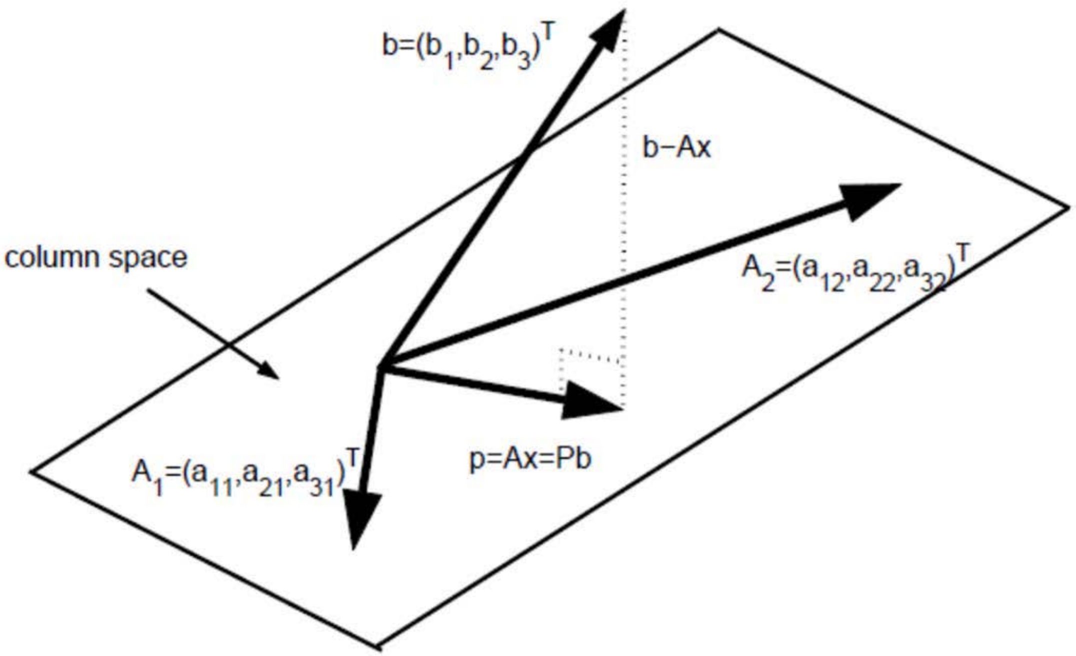
Least Squares

- $Ax = b$ 의 의미
  - 만약 $\boldsymbol{b} \notin \boldsymbol{R}(\boldsymbol{A})$라면?
  - $\boldsymbol{R}(\boldsymbol{A})$의 벡터 중 $\boldsymbol{b}$와 가장 가까운 벡터를 찾는 것
  - 가장 가까운 벡터?



Least Squares

- $Ax = b$ 의 의미
  - 만약 $\boldsymbol{b} \notin \boldsymbol{R}(\boldsymbol{A})$라면?
  - $\boldsymbol{R}(\boldsymbol{A})$의 벡터 중 $\boldsymbol{b}$와 가장 가까운 벡터를 찾는 것
  - 가장 가까운 벡터? = 수직으로 내린 벡터!



Least Squares

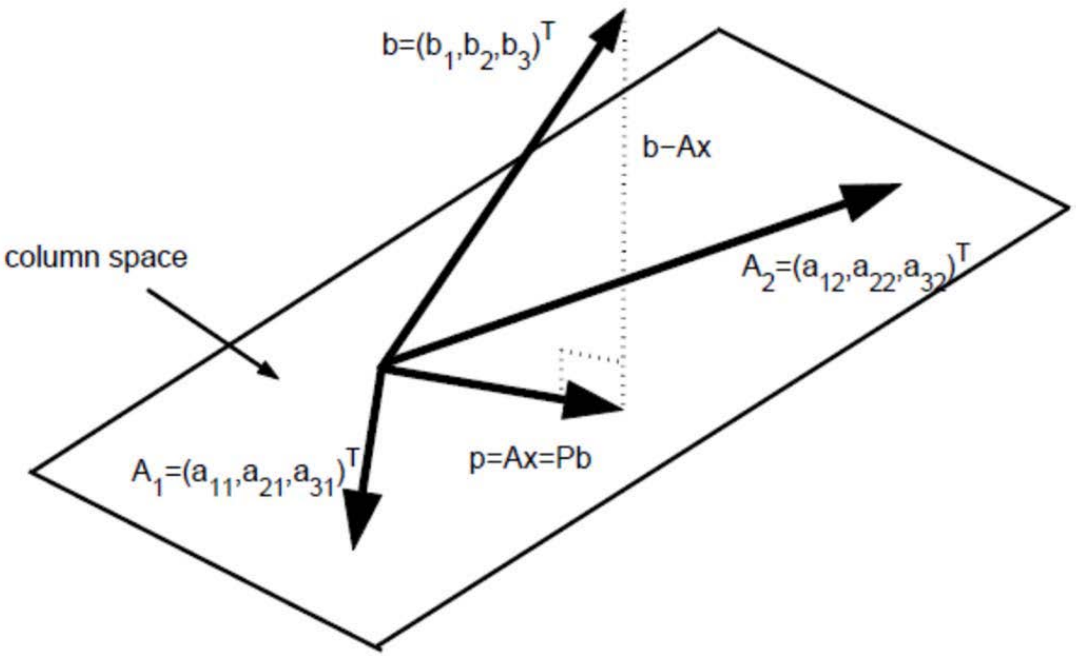
- *Orthogonality*
  - x, y가 다음과 같을 때 서로 수직(Orthogonal)하다고 말한다
  - $x^T y = x \cdot y = 0$
  - $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \end{bmatrix}$는 서로 수직이다
  - $\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 4 \end{bmatrix} = 1 \times 0 + 0 \times 4 = 0$

Least Squares

- *Orthogonality*
  - $A: \; m \times n \, matrix$
  - $\boldsymbol{R(A)}$ **: column space of** $\boldsymbol{A}$ = All linear combinations of the columns of A
  - $\boldsymbol{N(A)}$ **: null space of** $\boldsymbol{A}$ = All linear combinations of x such that $Ax = 0$
  - $\boldsymbol{R(A^T)}$ **: row space of** $\boldsymbol{A}$ = All linear combinations of the rows of A
  - $\boldsymbol{N(A^T)}$ **: left null space of** $\boldsymbol{A}$ = All linear combinations of y such that $A^T y = 0$
  - **R(A) 와 N($\mathbf{A^T}$) 는 수직**
  - N(A) 와 R($\mathrm{A^T}$) 는 수직

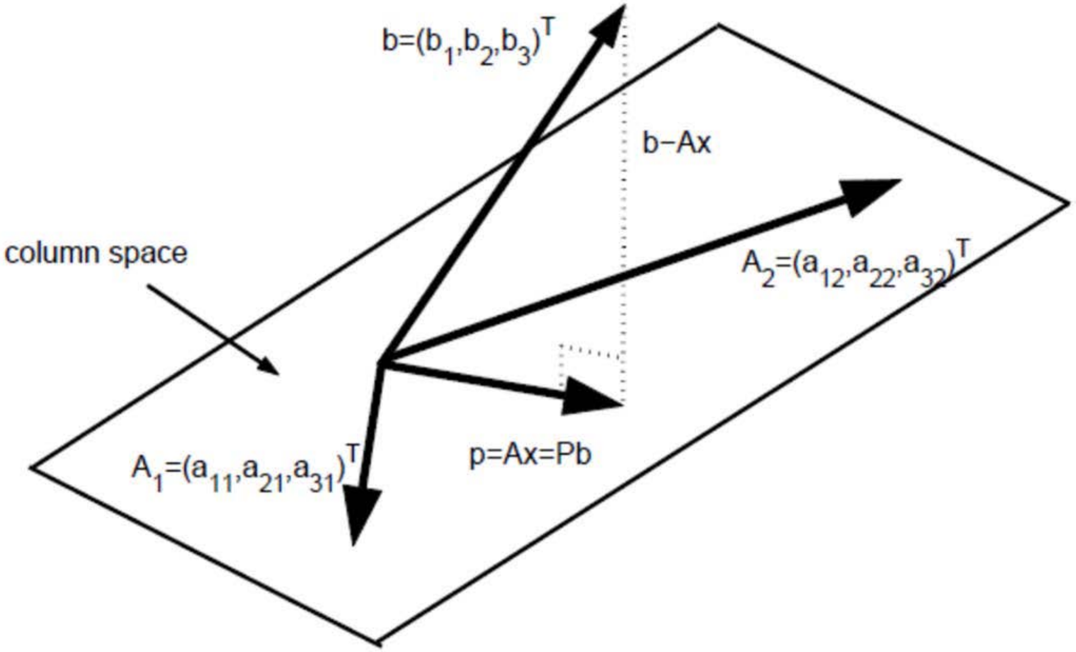
Least Squares

- Application of Fredholm alternative theorem
  - 임의의 벡터 x는 행렬 A에 대해 다음과 같이 수직 분해될 수 있다.
  - $x = x_r + x_n$
  - 단, $x_r \in R(A)$ & $x_n \in N(A^T)$



Least Squares

- Application of Fredholm alternative theorem
  - 임의의 벡터 x는 행렬 A에 대해 다음과 같이 수직 분해될 수 있다.
  - $x = x_r + x_n$
  - 단, $x_r \in R(A)$ & $x_n \in N(A^T)$
  - $x_n = Px = A(A^TA)^{-1}A^Tx$
  - $= orthogonal\ projection$



Least Squares

- Theorem (Least Square Solution)
  - $min||Ax - b||$
  - $p = A(A^T A)^{-1} A^T b = Ax$
  - $\hat{\boldsymbol{x}} = (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{A}^T \boldsymbol{b}$ ***: normal equations***