

데이터 벡터화

- 벡터화란 수학적인 의미로 행렬을 세로 벡터로 바꾸는 선형변환의 하나이다.

-

$$\text{행렬 } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$



$$\text{vec}(A) = \begin{bmatrix} a \\ c \\ b \\ d \end{bmatrix}$$

행렬 A를 벡터화하면 vec(A)가 된다.

머신러닝의 데이터

- 보통 matrix or table의 형태로 되어있다.
- 독립 변수(설명 변수)인 x와 종속 변수(반응 변수)인 y로 이루어져있다.
- 무작정 데이터를 수집하면 안된다. -> 데이터를 matrix 형태로 추출 후 벡터화하는 것이 좋다.
· *feature를 산출할 때, 최종 단계에서 vectorization을 실시함.*