

Statistic/Math

1. 고유값(eigen value)과 고유벡터(eigen vector)에 대해 설명해주세요. 그리고 왜 중요할까요?
2. 샘플링(Sampling)과 리샘플링(Resampling)에 대해 설명해주세요. 리샘플링은 무슨 장점이 있을까요?

Machine Learning

3. 정규화를 왜 해야할까요? 정규화의 방법은 무엇이 있나요?
4. Local Minima와 Global Minima에 대해 설명해주세요.
5. 차원의 저주에 대해 설명해주세요.

Deep Learning

6. 딥러닝은 무엇인가요? 딥러닝과 머신러닝의 차이는?
7. Cost Function과 Activation Function은 무엇인가요?
8. Data Normalization은 무엇이고 왜 필요한가요?

Python

9. What is the difference between list and tuples in Python?
10. What are the key features of Python?.

1. 정방행렬 A 에 대하여 $Ax = \lambda x$ 를 만족하는 λ 를 고유값, x 를 고유벡터라 한다.
 A 행렬에 대하여 고유벡터 x 는 좌표평면에서 방향을 의미하고, 고유값 λ 는 세기를 의미한다.

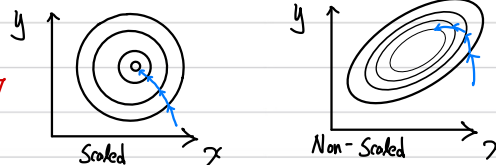
2. Sampling: 전수조사가 불가능한 모집단에 대하여 표본집단을 생성하는 행위

Resampling: 이미 뽑힌 표본집단에서 다시 부분 집합을 생성하여 분포를 관측하는 행위

→ 모집단의 어떤 가정도 필요 X

3. 정규화 이유: 빠른 학습 및 정확한 학습

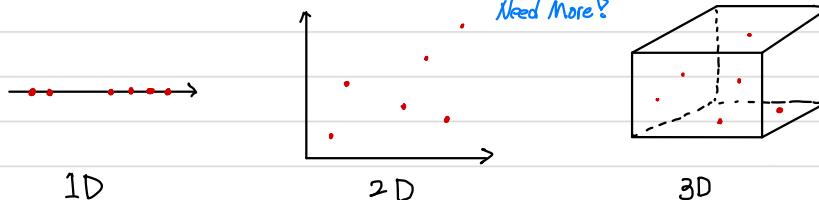
정규화 방법: Min-Max Scaler, Z-Score Scaler



4. Local Minima : 극소

Global Minima : 최소

5. 차원의 저주 : 데이터 학습을 위해 차원이 증가하면서 학습 데이터 수가 차원의 수보다 적어져 성능이 저하되는 현상

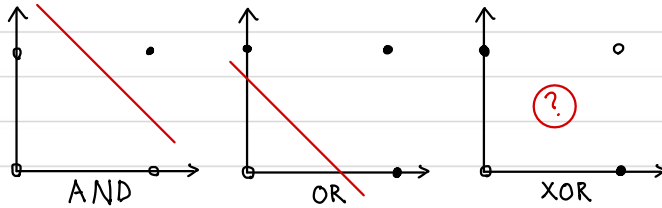


6. 인공지능: 인간의 지능으로 할 수 있는 사고, 학습 등을 컴퓨터가 할 수 있게 하는 기술
 머신러닝: 데이터를 기반으로 패턴을 학습하고 결과를 예측하는 알고리즘 기법
 딥러닝: 여러층을 인공신경망을 이용하여 머신러닝 기법을 수행하는 것

7. Cost Function: 모델의 실제값과 예측값과의 차이를 표현한 함수

Activation Function: 입력된 가중합을 출력 신호로 변환하는 함수. 주로 비선형 활성화 함수를 이용해 비선형 모델을 구성한다.

☆ 비선형 모델을 사용하는 이유는 데이터가 복잡해지고 차원이 커질수록 선형의 boundary로는 표현이 불가능하다.



8. Data Normalization: 데이터 스케일이 큰 특징이 더 강하게 반영되지 않도록 하기 위해 정규화를 진행한다.

→ 학습 속도 ↑, 정확도 ↑

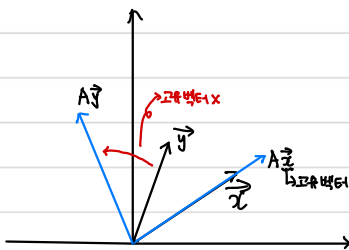
9. List: 속도가 느리지만, 변경 가능

Tuple: 속도가 빠르지만, 변경 불가능

10. Python의 특징?

- 대화형 인터프리터 언어 (사용하기 쉽고, 작성한 다음 바로 실행 가능)
- 동적인 데이터 타입
- 플랫폼 독립적 언어 - 대부분 운영 체제에서 동작
- 고수준의 내장 객체 자료형 제공
- 메모리 자동 관리
- 높은 확장성

본래 A 는 원차 변환을 시켜지만, A 호에서 방향이 바뀌지 않고 크기만 바뀔 경우, λ 를 고유 벡터라 한다.



즉, λ 은 실수배하는 역할이므로 $A' = \lambda A$ 라 표현할 수 있고, 이때 λ 를 고유값이라 한다.