

Review: Statistical & Geometric Perspective for Deep Learning

Ki Hyun Kim

nlp.with.deep.learning@gmail.com

Before this class,

- Our objective is:
 - 세상에 존재하는 어떤 미지의 함수를 모사하자.
- 주어진 입력(x)에 대해서 원하는 출력(y)을 반환하도록, 손실함수를 최소화하는 파라미터(θ)를 찾자.
- Gradient descent를 수행하기 위해 back-propagation을 수행하자.



After this class,

- Our objective becomes:
 - 세상에 존재하는 어떤 미지의 확률 분포 함수를 모사(approximate)하자.

After this class,

- Our objective becomes:
 - 세상에 존재하는 어떤 미지의 확률 분포 함수를 모사(approximate)하자.
- + Probabilistic Perspective
 - 확률 분포 $P(x)$ 와 $P(y|x)$ 로부터 데이터를 수집하여,
 - 해당 데이터를 가장 잘 설명하는 확률 분포 함수의 파라미터(θ)를 찾자: $\log P(y|x; \theta)$
 - Maximum Likelihood Estimation
 - Gradient Descent using Back-propagation
 - 또는 두 확률 분포를 비슷하게 만들자
 - Minimize Cross Entropy (or KL-Divergence)

After this class,

- Our objective becomes:
 - 세상에 존재하는 어떤 미지의 확률 분포 함수를 모사(approximate)하자.
- + Probabilistic Perspective
 - 확률 분포 $P(x)$ 와 $P(y|x)$ 로부터 데이터를 수집하여,
 - 해당 데이터를 가장 잘 설명하는 확률 분포 함수의 파라미터(θ)를 찾자: $\log P(y|x; \theta)$
 - Maximum Likelihood Estimation
 - Gradient Descent using Back-propagation
 - 또는 두 확률 분포를 비슷하게 만들자
 - Minimize Cross Entropy (or KL-Divergence)
- + Geometric Perspective
 - 데이터란 저차원의 manifold에 분포하고 있으며, 여기에 약간의 노이즈(ϵ)가 추가 되어 있는 것
 - 노이즈란 $\text{task}(x \rightarrow y)$ 에 따라서 다양하게 해석 가능 할 것
 - 따라서 해당 manifold를 배울 수 있다면, 더 낮은 차원으로 효율적인 맵핑(or project)이 가능
 - Non-linear dimension reduction

After this class,

- Our objective becomes:
 - 세상에 존재하는 어떤 미지의 확률 분포 함수를 모사(approximate)하자.
- + Probabilistic Perspective
 - 확률 분포 $P(x)$ 와 $P(y|x)$ 로부터 데이터를 수집하여,
 - 해당 데이터를 가장 잘 설명하는 확률 분포 함수의 파라미터(θ)를 찾자: $\log P(y|x; \theta)$
 - Maximum Likelihood Estimation
 - Gradient Descent using Back-propagation
 - 또는 두 확률 분포를 비슷하게 만들자
 - Minimize Cross Entropy (or KL-Divergence)
- + Geometric Perspective
 - 데이터란 저차원의 manifold에 분포하고 있으며, 여기에 약간의 노이즈(ϵ)가 추가 되어 있는 것
 - 노이즈란 task($x \rightarrow y$)에 따라서 다양하게 해석 가능 할 것
 - 따라서 해당 manifold를 배울 수 있다면, 더 낮은 차원으로 효율적인 맵핑(or project)이 가능
 - Non-linear dimension reduction
- + Representation Learning, Again
 - 낮은 차원에서의 표현을 통해,
차원의 저주(curse of dimensionality)를 벗어나 효과적인 학습이 가능

