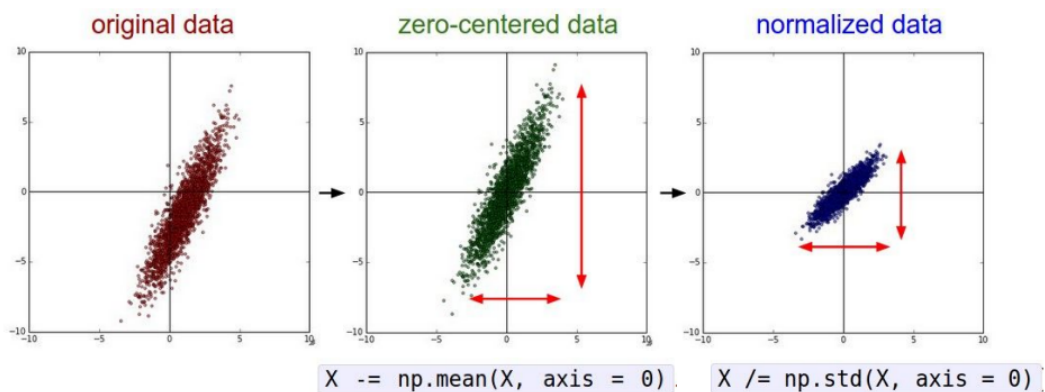


Lecture 6-2

Title	Training Neural Networks I
slide	http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture6.pdf

Data Preprocessing



네트워크를 훈련시키기 위해 우리에게 주어진 입력 데이터를 전처리를 해주어야 한다. 일반적인 전처리 과정은 "zero-mean"으로 만들고, "normalize" 하는 것이다. normalization은 보통 표준편차로 해준다. 입력이 모두 positive인 경우에는 모든 뉴런이 positive인 gradient를 얻게 되어 최적하지 못한(suboptimal) 최적화가 되게 되어, 이는 데이터가 모두 positive일 때 뿐만 아니라, 0이거나 모두 negative일 경우에도 발생한다. normalization을 해주는 이유는 모든 차원이 동일한 범위 안에 있게 해줌으로써 전부 동등한 기여(contribute)를 하게 한다.

TLDR: In practice for Images: center only

e.g. consider CIFAR-10 example with [32,32,3] images

- Subtract the mean image (e.g. AlexNet)
(mean image = [32,32,3] array)
- Subtract per-channel mean (e.g. VGGNet)
(mean along each channel = 3 numbers)

이미지의 경우에는 전처리로 zero-centering만 해주고, normalization는 하지 않는다. 왜냐하면 이미 이미지의 각 차원끼리 스케일이 어느 정도 맞춰져 있기 때문이다. 그래서 이미

지 데이터를 다룰 때는 단순히 zero-mean 정도만 사용하고, normalization와 그 밖의 여러 복잡한 방법들은 잘 사용하지 않는다. 또한 일반적으로 이미지를 다룰 때, 굳이 입력을 더 낮은 차원으로 projection 시키지 않고, CNN에서는 원본 이미지 자체의 spatial 정보를 이용해서 이미지의 spatial structure를 얻을 수 있도록 한다.

위의 그림에서 채널은 RGB이다. 예를 들어 이미지가 32x32x3라면 Width(32) x Height(32) x RGB(3) 가 된다. VGG식으로 하면 RGB별로 각각 평균을 구하는 것이다. 이때 평균은 "학습 데이터 전부" 를 가지고 계산한다. 만약 미니배치 단위로 학습시킨다고 하더라도 평균 계산은 미니배치 단위로 각각 하는 것이 아니라 전체로 계산한다. 배치에서 뽑은 데이터도 전체 데이터에서 나온 것들이므로 결론적으로 배치 평균과 전체 평균의 값은 (이상적으로) 같다. 따라서 처음에 전체 트레이닝 이미지를 이용하여 단 한번만 평균을 계산하는 것이 편하고, 만약 데이터가 엄청 크다면 모든 데이터를 전부 쓰지 않고 적절하게 샘플링해서 구할 수도 있다.