CS231n(3강)

최민우

- Linear Classifier

- f(x,W) = W*x + b
- 가중치(Weight)의 값이 랜덤
- 고양이, 자동차, 개구리 -> 개, 자동차, 트럭
- Loss(오차)가 발생

Recall from last time: Linear Classifier

airplane	-3.45	-0.51	3.42
automobile	-8.87	**6.04**	4.64
bird	0.09	5.31	2.65
cat	**2.9**	-4.22	5.1
deer	4.48	-4.19	2.64
dog	8.02	3.58	5.55
frog	3.78	4.49	**-4.34**
horse	1.06	-4.37	-1.5
ship	-0.36	-2.09	-4.79
truck	-0.72	-2.93	6.14

TODO:

1. Define a **loss function** that quantifies our unhappiness with the scores across the training data.
2. Come up with a way of efficiently finding the parameters that minimize the loss function. **(optimization)**

1.Loss Functions(손실함수)

- 알고리즘이 예측한 값과 실제 정답의 차이를 비교하기 위한 함수
- Optimization(최적화)을 위해 최소화하는 것이 목적

2.Multiclass SVM loss

Suppose: 3 training examples, 3 classes.
With some W the scores $f(x, W) = Wx$ are:

cat	**3.2**	1.3	2.2
car	5.1	**4.9**	2.5
frog	-1.7	2.0	**-3.1**
Losses:	2.9	0	12.9

Multiclass SVM loss:

Given an example (x_i, y_i)
where x_i is the image and
where y_i is the (integer) label,

and using the shorthand for the scores vector: $s = f(x_i, W)$

the SVM loss has the form:

$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + 1)$$

Loss over full dataset is average:

$$L = \frac{1}{N} \sum_{i=1}^{N} L_i$$

L = (2.9 + 0 + 12.9)/3
= **5.27**

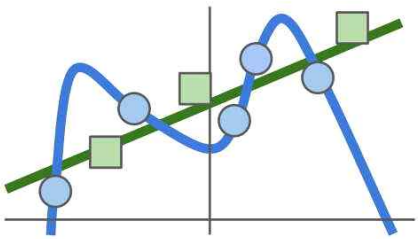
3.Regularization(λR(W))

- W가 너무 큰 값을 가지게 되면, 과하게 구불구불한 형태의 함수가 만들어 짐
- W(가중치)가 너무 큰 값들을 가지지 않도록 하기 위해서, 모델의 복잡도를 낮추어야 함
- L2 Regularization, L1 Regularization 등

$$L(W) = \underbrace{\frac{1}{N}\sum_{i=1}^{N} L_i(f(x_i, W), y_i)}_{\text{Data loss}} + \underbrace{\lambda R(W)}_{\text{Regularization}}$$

Data loss: Model predictions should match training data

Regularization: Model should be "simple", so it works on test data



Occam's Razor:
"Among competing hypotheses, the simplest is the best"
William of Ockham, 1285 - 1347

4.Softmax Classifier(확률)

Softmax Classifier (Multinomial Logistic Regression)

$$L_i = -\log\left(\frac{e^{sy_i}}{\sum_j e^{s_j}}\right)$$

unnormalized probabilities

	unnormalized log probabilities		probabilities	
cat	3.2	exp → 24.5	normalize → 0.13	→ L_i = -log(0.13) = 0.89
car	5.1	164.0	0.87	
frog	-1.7	0.18	0.00	

5.Optimization(최적화)

- Loss Function(손실함수) 값을 최소화하는 파라미터를 구하는 과정 : 경사하강법