

Machine Learning to Deep Learning

Palash Chauhan

May 9, 2017

1 Multinomial Logistic Classification

1.1 Softmax

- Scores \rightarrow probabilities
- Multiply by 10 \rightarrow close to 0/1
- Divide by 10 \rightarrow close to uniform

1.2 Cross Entropy

- $D(S, L) = -\sum_i L_i \log(S_i)$
- L are true one hot labels, S are output of softmax from the model
- Minimize average cross entropy (loss) w.r.t parameters and biases to learn

1.3 Numerical Stability

- Loss function should never get too big or too small
- We want variables to always have 0 mean and equal variances
- For images (0-255), subtract 128 and divide by 128
- **Initialization:** Draw weights and biases from a gaussian with mean μ and small variance σ .

1.4 Measuring Performance

- Train, Test, Validation
- Use a validation set to prevent overfitting on test set
- A change that affects 30 examples in the validation set is significant and can be trusted