

# Progress report

Ching-Yuan, Tsai

2017,11,30  
Presented by Ching-Yuan, Tsai

# Outline

- 1 Bottleneck
- 2 Weights distribution
- 3 Threshold trend
- 4 Adjust threshold

Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

Network bandwidth  
Prune weights(gradients)

# Network bandwidth

- 1G bits/s.
- 32M floats/s.

## Prune by threshold

- One standard deviation.
- Two standard deviation.
- Dynamic adjust threshold.

# Outline

- 1 Bottleneck
- 2 Weights distribution
- 3 Threshold trend
- 4 Adjust threshold

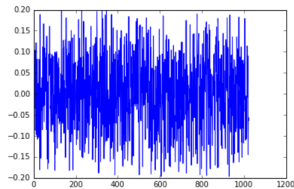
# Mnist handwritten database

- Two convolutional layers.
- Two max pooling layers.
- Two dense layers.
- One dropout layer.

Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

Model  
**Initial**  
Iteration 50  
Iteration 200  
Iteration 550  
Iteration 950

# Initial

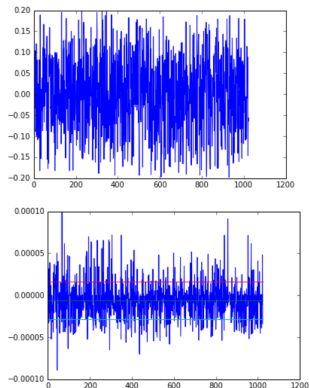


initial

Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

Model  
Initial  
Iteration 50  
Iteration 200  
Iteration 550  
Iteration 950

# Weights and gradients

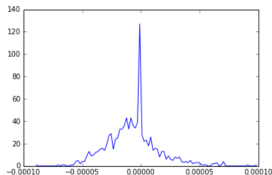
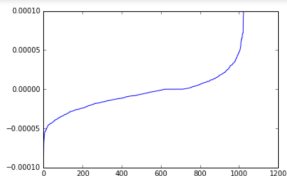




Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

Model  
Initial  
Iteration 50  
Iteration 200  
Iteration 550  
Iteration 950

# Gradients distribution

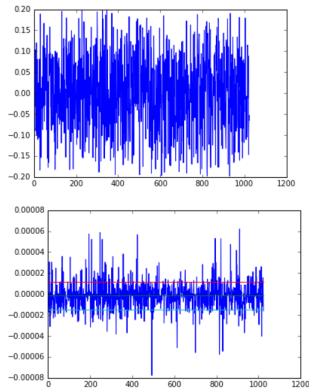


step 51, training accuracy 0.84

Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

Model  
Initial  
Iteration 50  
**Iteration 200**  
Iteration 550  
Iteration 950

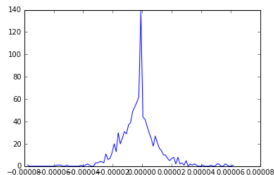
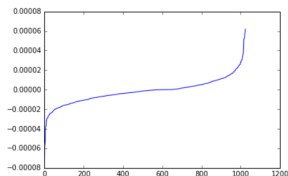
# Weights and gradients



Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

Model  
Initial  
Iteration 50  
**Iteration 200**  
Iteration 550  
Iteration 950

# Gradients distribution

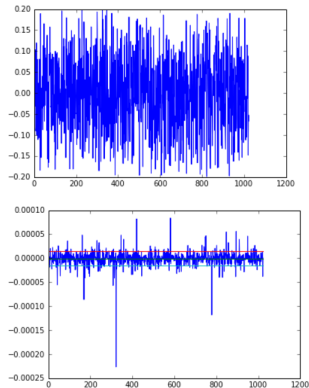


step 201, training accuracy 0.8

Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

Model  
Initial  
Iteration 50  
Iteration 200  
Iteration 550  
Iteration 950

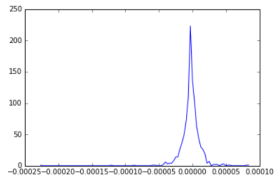
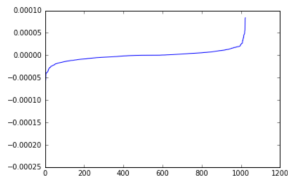
# Weights and gradients



Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

Model  
Initial  
Iteration 50  
Iteration 200  
**Iteration 550**  
Iteration 950

## Gradients distribution

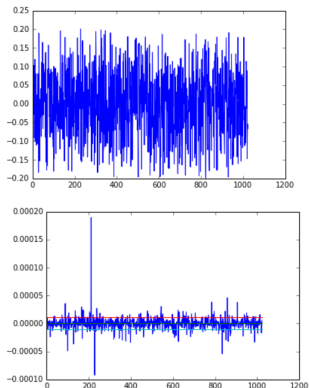


step 551, training accuracy 0.96

Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

Model  
Initial  
Iteration 50  
Iteration 200  
Iteration 550  
Iteration 950

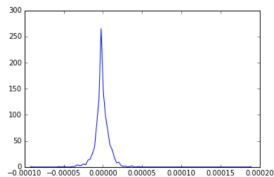
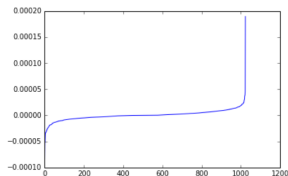
# Weights and gradients



Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

Model  
Initial  
Iteration 50  
Iteration 200  
Iteration 550  
Iteration 950

# Gradients distribution



step 951, training accuracy 1  
test accuracy 0.9635

# Outline

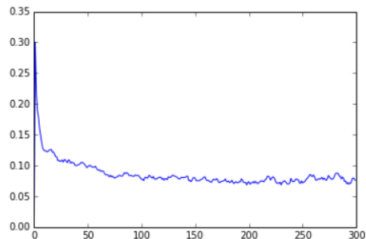
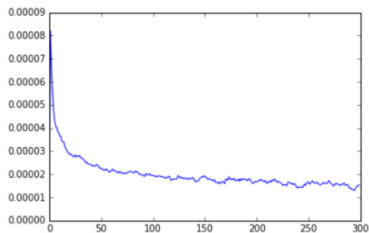
- 1 Bottleneck
- 2 Weights distribution
- 3 Threshold trend**
- 4 Adjust threshold



Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

One standard deviation  
Two standard deviation  
Accuracy

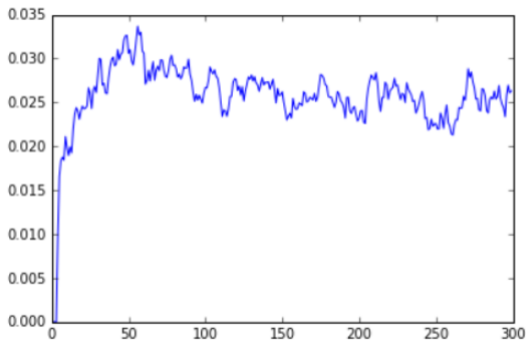
## One standard deviation



Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

One standard deviation  
Two standard deviation  
Accuracy

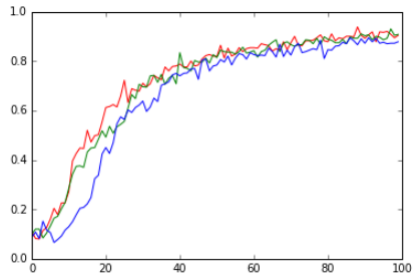
## Two standard deviation



Bottleneck  
Weights distribution  
Threshold trend  
Adjust threshold

One standard deviation  
Two standard deviation  
Accuracy

# Accuracy



# Outline

- 1 Bottleneck
- 2 Weights distribution
- 3 Threshold trend
- 4 Adjust threshold

# Adjust threshold

- New threshold = mean of the gradients which are not sent.
- Accumulate gradients.