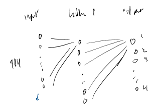
Nn math

Sunday, April 13, 2025 10:34 PM

Grad descent:
- Loop:
  - Forward
  - Backward
  - Update

Forward pass



0–255

[0, 1, 4, 7, 233, …, 0, 1]

each col is a new image

m

20 × 784 × 784 × m + 20 × m ⇒ 20 × m

weights for Im1

weights for Im2

20 × 784 × 784 × m

Im1 Im2

$$z^{[1]} = w^{[1]} \cdot A^{[0]} + b^{[1]}$$

$$A^{[1]} = \text{relu}(z^{[1]})$$

↳ np.max (matrix, 0)

10 × 20 · 20 × m + 10 × m

$$z^{[2]} = w^{[2]} \cdot A^{[1]} + b^{[2]}$$

$$A^{[2]} = \text{softmax}(z^{[2]})$$

↳ $\dfrac{\text{np.exp(matrix)}}{\text{np.sum(np.exp(matrix))}}$

does np.sum $A^{[2]}$ = m or 1

good / problem

$z^{[2]}$ has dim. 10×m

want to softmax along every col

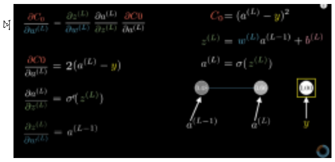
☆ init dimensionality

☆ make ff_net

☆ ensure sum of matrix is good

backward math

how much to change 2nd layer given the loss = $dw^{[2]}$



The loss

$$dz^{[2]} = A^{[2]} - Y$$

$$\frac{\partial a^{[2]}}{\partial z^{[2]}} = \sigma'(z^{[2]})$$

↓ softmax ↓ baked in here? ↓

How much each activation in a[1] contributed to the error
Divided by m to average across all training samples

$$dW^{[2]} = \frac{1}{m}(A^{[2]} - Y) \cdot A^{[1]T}$$

10 × m   m × 20

How much each bias should change on average across all
Samples, this is just a flat number so using just A[2]-Y is sufficient
Don't need to do it wrt. to any other A[1]

$$db^{[2]} = \frac{1}{m}\sum (A^{[2]} - Y)$$

Takes W1 dot first image cost , when W1 dot second image cost -
Lets us associate which weights responsible for the output error

$$dz^{[1]} = W^{[2]T} dz^{[2]} \cdot \text{relu}'(z^{[1]})$$

20 × 10   10 × m

Largely same as above first two

$$dw^{[1]} = \frac{1}{m} dz^{[1]} \cdot A^{[0]T}$$

$$db^{[1]} = \frac{1}{m}\sum dz^{[1]}$$

☆ loss calculation : cross entropy loss

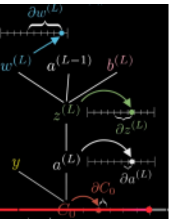
☆ one hot encode (Y):

init zeros in shape m × 10

Y = [1, 4, 7, 2 … , 2, 1] of length m

10 × m matrix^T = m × 10 matrix

for i in Y,
into temp [i, Y[i]] = 1



Update

$w^{[1]} := w^{[1]} - \alpha\, dw^{[1]}$

$b^{[1]} := b^{[1]} - \alpha\, db^{[1]}$

$w^{[2]} := w^{[2]} - \alpha\, dw^{[2]}$

$b^{[2]} := b^{[2]} - \alpha\, db^{[2]}$