Tuesday, Feb 7

Neural Networks
SpeakUp

Join by URL: https://web.speakup.info/room/join/80740

Join by key: web.speakup.info → use key 80740

Join by saved session: web.speakup.info

Join by QR:
Outline

• Questions about projects?
• Recap of a singular neuron model
• Notation and terminology
• Compute graphs
• Optimization
• Backpropagation

• This will be our most math heavy week

• Next week we’ll rely on PyTorch to compute all derivatives
Recap: A Single Neuron

• Take five minutes to draw
  • Whatever will help you remember (no correct or incorrect drawings)
  • You’ll keep a running drawing log the rest of the semester
Fully-Connected (Feed-Forward) Network

\[ Z = \text{input} - \text{Doram} + \text{poram} \]

\[ \alpha = g(z) \]

Activation:

\[ z_j = \sum_{i=0}^{n} w_{j,i} \cdot a_i + b \]

Linear:

\[ a_j = g(z_j) \]
Fully-Connected (Feed-Forward) Network
Fully-Connected (Feed-Forward) Network

Simplified (easier to draw) diagram
Vectorized Equations

\[
Z^{[l+1]} = A^{[l]} W^{T} + b^{[l]}
\]

\begin{align*}
&= (N, n_{l-1}) (n_{l-1}, n_{l}) + (1, n_{l}) \\
&\quad \text{broadcasting} \\
&= (N, n_{l}) + (N, n_{l}) \\
&\quad \text{element-wise application}
\end{align*}

\[
z_{i}^{[l+1]} = \sum_{i=0}^{n_{l-1}} w_{j,i}^{[l]} a_{i}^{[l]}
\]

\[
a_{j}^{[l+1]} = g \left( z_{j}^{[l+1]} \right)
\]
Data

\[ N \rightarrow \text{# of training examples} \]

\[ N_x \rightarrow \text{# of features} \]

\[ y^{(3)} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \]

\[ y = \begin{bmatrix} y_{1(n)} \\ \vdots \\ y_{N(n)} \end{bmatrix} \]

\[ x^{(57,122)} = \begin{bmatrix} x_{1} \\ \vdots \\ x_{N} \end{bmatrix} \]

\[ x = \begin{bmatrix} x_{1} \\ \vdots \\ x_{N} \end{bmatrix} \]
MNIST Dataset Example

- MNIST includes 60,000 training images
- Each image is grayscale and 28x28 pixels in size
- Each output is a one-hot encoding of the digits 0 through 9

What is the shape of $X$?

$$ (N, n_x) \rightarrow (60,000, 784) $$

What is the shape of $Y$?

$$ (N, n_y) \rightarrow (60,000, 10) $$
MNIST Neural Network

- Imagine we have a two-layer network
- The hidden layer has 17 neurons

- What is the shape of $W^{[1]}$?
  $(17, 784)$

- What is the shape of $b^{[1]}$?
  $(17, 1)$

- What is the shape of $W^{[2]}$?
  $(10, 17)$

- What is the shape of $b^{[2]}$?
  $(10, 1)$
Vectorized Equations

• What is the shape of $Z^{[1]}$?
  $$(60,000, 17)$$

• What is the shape of $A^{[1]}$?
  $$(60,000, 17)$$

• What is the shape of $Z^{[2]}$?
  $$(60,000, 10)$$

• What is the shape of $A^{[2]}$?
  $$(60,000, 10)$$
Compute Graph

\[
Z = A_0 \cdot W + b
\]

\[
A_1 = g(Z)
\]

\[
Z = A_1 \cdot W + b
\]

\[
A = g(Z)
\]
Optimization with Half MSE and Sigmoid

\[ L = -|| y \log \hat{y} + (1-y) \log (1-\hat{y}) || \]

\[ g(z_{\text{out}}) = \sigma(g(z)) \]

\[ \frac{\partial}{\partial z} \sigma(z) = \sigma(z)(1 - \sigma(z)) \]
\[ \hat{y} = [0, 1] \]

\[ y \in \{0, 1\} \]

\[ L = -\sum \tag{AC03} \left( y \log \hat{y} + (1 - y) \log (1 - \hat{y}) \right) \]

| \( \hat{y} \) | \( y \) | \( \log \hat{y} \) | \( \log(1 - \hat{y}) \) | \( L \) |
|-------|-----|orable|orable| 0.1 |
| 0.1   | 0   | -2.3 | 0.1 | 0.1 |
| 0.1   | 1   | -2.3 | 0.1 | 2.3 |
| 0.9   | 0   | -0.1 | -2.3 | 2.3 |
| 0.9   | 1   | -0.1 | -2.3 | 0.1 |
Backpropagation $W^{[2]}$

\[
\frac{\partial L}{\partial W^{[2]}} = \frac{\partial}{\partial W^{[2]}} \left( \| \nabla \log \hat{y} + (1 - \hat{y}) \log (1 - \hat{y}) \| \right)
\]

1. \[\frac{\partial L}{\partial \hat{y}} = - \left( \frac{\hat{y}}{\hat{y}} - \frac{(1 - \hat{y})}{(1 - \hat{y})} \right)\]

2. \[\frac{\partial \hat{y}}{\partial z^{[2]}} = \sigma(z^{[2]}) (1 - \sigma(z^{[2]})) = \hat{y} \cdot (1 - \hat{y})\]

3. \[\frac{\partial z^{[2]}}{\partial W^{[2]}} = A^{[1]}\]

\[
\frac{\partial L}{\partial W^{[2]}} = - \left( \frac{\hat{y}}{\hat{y}} - \frac{(1 - \hat{y})}{(1 - \hat{y})} \right) \cdot \hat{y} \cdot (1 - \hat{y}) A^{[2]} = (\hat{y} - \hat{y}) A^{[2]}\]
Backpropagation $b^{[2]}$

\[
\frac{\partial L}{\partial b^{[2]}} = \frac{\partial}{\partial b^{[2]}} \left[ -\left( \frac{1}{\hat{y}} - \frac{1}{1-\hat{y}} \right) \right] \\
\frac{\partial L}{\partial b^{[2]}} = \frac{\partial}{\partial \hat{y}} \left( \frac{1}{\hat{y}} - \frac{1}{1-\hat{y}} \right) \times \frac{\partial \hat{y}}{\partial b^{[2]}}
\]

\[
\frac{\partial L}{\partial \hat{y}} = \frac{1}{\hat{y}} - \frac{1}{1-\hat{y}} \\
\frac{\partial \hat{y}}{\partial z^{[2]}} = \sigma(z^{[2]}) (1 - \sigma(z^{[2]})) \\
\frac{\partial z^{[2]}}{\partial b^{[2]}} = 1
\]
Backpropagation $W^{[1]}$

$$\frac{\partial L}{\partial W^{[2]}} = \frac{\partial L}{\partial y^{[3]}} \cdot \frac{\partial y^{[3]}}{\partial z^{[2]}} \cdot \frac{\partial z^{[2]}}{\partial A^{[2]}} \cdot \frac{\partial A^{[2]}}{\partial z^{[2]}} \cdot \frac{\partial z^{[2]}}{\partial W^{[2]}}$$
Backpropagation $b^{[1]}$
Parameter Updates

For Loop:

\[ w_{z2}^{[2]} = \alpha \cdot (\hat{y} - y) \cdot \delta_{z1} \]

\[ b_{z2} = \alpha \cdot (\hat{y} - y) \]

\[ w_{17} = \alpha \]

\[ b_{17} = \alpha \]