Lecture 2
Linear Regression

Sung Kim <hunkim+ml@gmail.com>
Acknowledgement

• Andrew Ng’s ML class
  - https://class.coursera.org/ml-003/lecture
  - http://www.holehouse.org/mlclass/ (note)

• Convolutional Neural Networks for Visual Recognition.
  - http://cs231n.github.io/

• Tensorflow
  - https://www.tensorflow.org
  - https://github.com/aymericdamien/TensorFlow-Examples
Predicting exam score: regression

<table>
<thead>
<tr>
<th>x (hours)</th>
<th>y (score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>9</td>
<td>80</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
</tr>
</tbody>
</table>
Regression (data)

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
Regression (presentation)

<table>
<thead>
<tr>
<th>x</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
(Linear) Hypothesis

![Graph showing linear hypothesis with data points]

- X-axis: 0 to 3
- Y-axis: 0 to 3
- Data points at (0,1), (1,2), (2,3), (3,3)
(Linear) Hypothesis
(Linear) Hypothesis

\[ H(x) = Wx + b \]
Which hypothesis is better?
Which hypothesis is better?
Cost function

- How fit the line to our (training) data

\[ H(x) - y \]
Cost function

• How fit the line to our (training) data

\[
\text{cost} = \frac{1}{m} \sum_{i=1}^{m} (H(x^{(i)}) - y^{(i)})^2
\]

\[
H(x) = Wx + b
\]
Cost function

\[
\text{cost} = \frac{1}{m} \sum_{i=1}^{m} (H(x^{(i)}) - y^{(i)})^2
\]

\[H(x) = Wx + b\]

\[
\text{cost}(W, b) = \frac{1}{m} \sum_{i=1}^{m} (H(x^{(i)}) - y^{(i)})^2
\]
Goal: Minimize cost

\[
\minimize_{W,b} \text{cost}(W, b)
\]