Lecture 01

Tom Cheung

October 2018

Supervised learning

feature	**target**
$x^{(0)}$	$y^{(0)}$
$x^{(1)}$	$y^{(1)}$
$\vdots$	$\vdots$
$x^{(i)}$	$y^{(i)}$

- training example: $(x^{(i)},\ y^{(i)})$
- training set: $\left\{(x^{(i)},\ y^{(i)}); i = 1, \cdots, m\right\}$
- space of input and output $\mathcal{X}, \mathcal{Y}$
- hypothesis: $h : \mathcal{X} \mapsto \mathcal{Y}$

REGRESSION *continuous*

CLASSIFICATION *discrete*

1 Linear Regression

feature selection*!?* — stay tuned

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 \tag{1}$$

θ_i : parameter or weight $\qquad$ $x_0 = 1$: intercept term

$$h(x) = \sum_{i=0}^{n} \theta_i x_i = \theta^{\mathrm{T}} x \tag{2}$$

cost function

$$J(\theta) = \frac{1}{2} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2 \tag{3}$$

The ordinary least squares regression model ***!?*** [1]

[1] The method of ordinary least squares can be used to find an approximate solution to overdetermined systems.
see ☞: **Approximate solutions**