Lecture14

Tuesday, October 30, 2018 9:28 AM

STATISTICAL MACHINE LEARNING

- Given a data set $S := \{ \vec{x}^i, y^i \}_{i=1}^N$

$\vec{x}^i \in X \subset \mathbb{R}^k$

↳ closed and bounded (compact)

$y^i \in Y \subset \mathbb{R}$

Def (ML in this class): ML is the task of learning the unknown function $f(\vec{x}): y^i = f(\vec{x}^i)$ given an incomplete information described by the data set $S := \{ \vec{x}^i, y^i \}_{i=1}^N$

- Define the product space $Z := X \times Y$, equipped with a "Borel" probability measure $\pi(\vec{x}, y)$. Let $\pi(y|\vec{x})$ be the conditional probability measure of y given $\vec{x}$ and

$$\pi(\vec{x}) := \int_Y \pi(\vec{x}, y)\, dy = \int_Y \pi(\vec{x}, dy)$$

Π by def. Bayes' theorem

$$\Pi(\vec{x}, y) \overset{??}{=} \Pi(y \mid \vec{x}) \times \Pi(\vec{x})$$

- Least square error (risk) w.r.t. $\Pi(\vec{x}, y)$

$$R(h) := \int_Z (h(\vec{x}) - y)^2 \, d\Pi(x, y)$$

where $h(\vec{x})$ is an approximation of $y = f(\vec{x})$