

# Introduction to Statistical Machine Learning

## Homework 1

Yota Toyama

October 19, 2016

1.

$$\begin{aligned}
 \frac{\partial}{\partial \mathbf{w}} R(\mathbf{w}) &= 0 \\
 \frac{\partial}{\partial \mathbf{w}} E_{p(\mathbf{x}, y)} [(y - \mathbf{w}^T \mathbf{x})^2] &= 0 \\
 E_{p(\mathbf{x}, y)} [2(y - \mathbf{w}^T \mathbf{x})(-\mathbf{x})] &= 0 \\
 E_{p(\mathbf{x}, y)} [(y - \mathbf{w}^T \mathbf{x})\mathbf{x}] &= 0 \\
 \mathbf{a}^T E_{p(\mathbf{x}, y)} [(y - \mathbf{w}^T \mathbf{x})\mathbf{x}] &= 0 \\
 E_{p(\mathbf{x}, y)} [(y - \mathbf{w}^T \mathbf{x})\mathbf{a}^T \mathbf{x}] &= 0
 \end{aligned}$$

2.

$$\begin{aligned}
 E_{p(\mathbf{x}, y)} [y - \mathbf{w}^T \mathbf{x}] &= 0 \\
 \therefore E_{p(\mathbf{x}, y)} [(y - \mathbf{w}^T \mathbf{x})E_{p(\mathbf{x})} [\mathbf{x}]] &= 0 \\
 E_{p(\mathbf{x}, y)} [(y - \mathbf{w}^T \mathbf{x})\mathbf{x}] + E_{p(\mathbf{x}, y)} [(y - \mathbf{w}^T \mathbf{x})E_{p(\mathbf{x})} [\mathbf{x}]] &= 0 \\
 E_{p(\mathbf{x}, y)} [(y - \mathbf{w}^T \mathbf{x})(\mathbf{x} - E_{p(\mathbf{x})} [\mathbf{x}])] &= 0
 \end{aligned}$$

$\therefore$  the correlation between data and prediction errors is 0.

3.