Introduction to Statistical Machine Learning
Homework 1

Yota Toyama

October 19, 2016

1.

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

2.

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

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

3.