# Exercises for Lecture 2

## 1.(Bishop, exercise 1.1)

Remember the polynomial curve fitting example. Given a set of $N$ training samples, $\{x, t\}_{n=1}^{N}$, we assume the following functional form for y.

$$y(x, \mathbf{w}) = \sum_{m=1}^{M} w_m x^m$$

Then we minimize the following sum-of-squares error loss.

$$E(w) = \frac{1}{2} \sum_{n=1}^{N} (y(x_n, \mathbf{w}) - t_n)^2$$

Show that the $\mathbf{w} = \{w_i\}$ that minimizes the above sum-of-squares loss are given by the solutions to the following set of linear equations.

$$\sum_{j=0}^{M} A_{ij} w_j = T_i$$

where

$$A_{ij} = \sum_{n=1}^{N} (x_n)^{i+j} \qquad T_i = \sum_{n=1}^{N} (x_n)^i t_n$$

## 2. Bias-variance decomposition

Prove the following.

$$\mathbb{E}_D[(y(x, D) - t)^2] = (\mathbb{E}_D[y(x, D)] - t)^2 + \mathbb{E}_D[(y(x, D) - \mathbb{E}_D[y(x, D)])^2]$$

You may find the following useful

- $\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$ for $x, y$ independent.
- $\mathbb{E}[kx] = k\mathbb{E}[x]$ for constant $k$.