Exercises for Lecture 7

1. Binary logistic regression

Consider the logistic regression problem with 2 classes where x denotes a training sample, t its label with the prediction from logistic regression model given as the following

$$y = \sigma(w^T x), \quad \sigma(a) = \frac{1}{1 + \exp(-a)}$$

where w are the parameters.

Cross-entropy error function for a single training sample x can be written as

$$E(w) = -t \ln y - (1 - t) \ln(1 - y)$$

Show that the gradient of the error function with respect to parameters is given by

$$\nabla_w E(w) = (y - t)x$$

2. Maximum likelihood for linear regression

Consider the linear regression model from a probabilistic perspective where we assume the target t for a given sample x has the following Normal distribution.

$$t \sim \mathcal{N}(w^T x, \sigma^2)$$

where w are the parameters (and σ^2 is a fixed variance).

Show that maximizing the likelihood $p(t|w, x)$ with respect to w is equivalent to minimizing the following sum-of-squares error function

$$E(w) = (t - w^T x)^2$$