Homework #1

Due on February 11th, 2020
Machine Learning
CS559WS−Spring 2020
Professor In Suk Jang

Daniel Kadyrov
10455680

Problem 1

By using a change of variables, verify that the univariate Gaussian distribution given by:

$$N(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{1}{2\sigma^2}(x-\mu)^2\}$$

satisfies $E(x) = \mu$. Next, by differentiating both sides of normalization condition

$$\int_{-\infty}^{-\infty} N(x \mid \mu, \sigma^2) dx = 1$$

with respect to σ^2, verify that the Gaussian satisfies $E(x^2) = \mu^2 + \sigma^2$.

Problem 2

Use $E(x) = \mu$ to prove $E(xx^T) = \mu\mu^T + \Sigma$. Now, using the results two definitions, show that:

$$E[x_n x_m] = \mu\mu^T + I_{nm}\Sigma$$

where x_n denotes a data point same from a Gaussian distribution with mean μ and covariance Σ,and I_{nm} denotes the (n, m) element of the identity matrix. Hence prove the result (2.124)

Problem 3

Consider a linear model of the form:

$$f(x, w) = w_0 + \sum_{i=1}^{D} w_i x_i$$

together with a sum of squares/loss function of the form:

$$L_D(w) = \frac{1}{2} \sum_{n=1}^{N} (f(x_n, w) - y_n)^2$$

Now suppose that Gaussian noise ϵ_i with zero mean and variance σ^2 is added independently to each of the input variables x_i. By making use of $E[\epsilon_i] = 0$ and $E[\epsilon_i \epsilon_j] = \delta_{ij}\sigma^2$, show that minimizing L_D averaged over the noise distribution is equivalent to minimizing the sum of square error for noise-free input variables with the addition of a weight-decay regularization term, in which the bias parameters w_0 is omitted from the regularizer.