Inleiding Machine Learning

Exercises, week 2

Exercise 1

Bishop ex. 1.4 (WWW)

Exercise 2

Probability densities $p(x)$ should be non-negative $p(x) \geq 0$, and normalised $\int p(x)dx = 1$.

1. Consider the probability density $p(t)$ defined as

$$p(t) = \begin{cases} \frac{1}{Z}\exp(-\lambda t) & , \quad t \geq 0 \\ 0 & , \quad t < 0 \end{cases} \tag{1}$$

   with λ a positive constant. Compute Z using the fact that p should be normalised.

2. Let $\rho(x)$ be a normalised probability density, i.e. $\rho(x) >= 0$ and $\int_{-\infty}^{\infty} \rho(x)dx = 1$. Show that for any pair of constants μ and $\alpha > 0$, the function

$$\hat{\rho}(x) = \alpha\, \rho(\alpha(x - \mu)) \tag{2}$$

   is also a normalised density.

3. Compute the normalising constant Z of the following probability density in R^d with parameters $\lambda_i > 0$,

$$p(x_1, \ldots, x_d) = \frac{1}{Z}\exp\left\{-\sum_{i=1}^{d} \frac{\lambda_i}{2} x_i^2\right\} . \tag{3}$$

   You may use that for $\lambda > 0$,

$$\int_{-\infty}^{\infty} \exp\left\{-\frac{\lambda}{2}x^2\right\} dx = \left(\frac{2\pi}{\lambda}\right)^{1/2}$$

Exercise 3

The variance of f is defined as

$$\mathrm{var}[f] = \left\langle (f(x) - \langle f(x) \rangle)^2 \right\rangle \tag{4}$$

in which $\langle f(x) \rangle \equiv \mathbb{E}[f]$ is the expectation of a function $f(x)$ under probability distribution $p(x)$, defined as $\mathbb{E}[f] = \int f(x)p(x)\, dx$. Now show that the variance can also be written as

$$\mathrm{var}[f] = \left\langle f(x)^2 \right\rangle - \left\langle f(x) \right\rangle^2 \tag{5}$$

Exercise 4 is optional, recommended for those who are unfamiliar with expectations and variances

Exercise 4

More about expectation values and variances.

Consider a discrete random variable x with distribution $p(x)$. The expectation of a function $f(x)$ is

$$\mathbb{E}[f] = \sum_x p(x)f(x) \tag{6}$$

Its variance $\mathrm{var}[f]$ is

$$\mathrm{var}[f] = \mathbb{E}[f^2] - (\mathbb{E}[f])^2 \tag{7}$$

- Show that if c is a constant,

$$\begin{aligned} \mathbb{E}[cf] &= c\mathbb{E}[f] && (8) \\ \mathrm{var}[cf] &= c^2\mathrm{var}[f] && (9) \end{aligned}$$

We now consider two discrete random variables x and z with a joint probability distribution $p(x, z)$. The expectation of a function $f(x, z)$ of x and z is given by

$$\mathbb{E}[f] = \sum_{x,z} p(x, z)f(x, z) \tag{10}$$

1. Show, using (10) that the expectation of the sum of x and z satisfies

$$\mathbb{E}[x + z] = \mathbb{E}[x] + \mathbb{E}[z] \tag{11}$$

   (Hints: make use of marginal distributions $p(z) = \sum_x p(x, z)$.)

2. Show that if x and z are statistical independent, i.e.,$p(x, z) = p(x)p(z)$, the expectation of their product satisfies

$$\mathbb{E}[xz] = \mathbb{E}[x]\mathbb{E}[z] \tag{12}$$

3. Use (7) and results (11) and (12) to show that the variance of the sum of two independent variables x and z satisfies

$$\mathrm{var}[x + z] = \mathrm{var}[x] + \mathrm{var}[z] \tag{13}$$

   (Hint: use that square of any sum $a + b$ satisfies $(a + b)^2 = a^2 + 2ab + b^2$)

Note: the properties of expectations and variance that are shown in this exercise hold for continuous variables as well, this can be shown in a similar way (i.e. by replacing sums by integrals.)

Exercise 5

Given the one dimensional Gaussian distribution Eq.1.46.

- show that the distribution is normalized
- show that the expected x is μ
- show that the expected variance is σ^2.

Exercise 6

Bishop ex. 1.11

Exercise 7

Show that the maximum likelihood estimator σ^2_{ML} in Eq. 1.58 is biased.

Exercise 8

Bishop ex. 1.17 (WWW)

Exercise 9

Bishop ex. 1.18 (WWW)

Exercise 10

Bishop ex. 1.19