IFT 6390 Homework 1

Author

September 20, 2018

Problem 1. Small exercise on probabilities

Proof. p ☐

Problem 3.1 (a). Name parameters

Proof. the parameters are σ and μ
$\mu \in (d, 1)$
The matrix Σ is d by d, where the diagonal terms are σ ☐

Problem 3.1 (b). Equation for optimal parameters

Proof.

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\sigma = \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)^T$$

☐

Problem 3.1 (c). What is the algorithmic complexity of this training method, i.e. of the method calculating these parameters?

Proof. To computer μ:
for i = 1 to n
sum each x_i with d components
endfor
Total is $O(nd)$

To computer σ:
for i = 1 to n
subtract μ from each x_i with d components
repeat the last operation with transpose multiply d by 1 and 1 by d vectors, result in a d by d matrix, which is d^2 operations endfor
Total is $O(nd^2 + d + d) = O(nd^2)$

☐

Problem 3.4 (a). Express the equation of a diagonal Gaussian density in Rd. Specify what are its parameters and their dimensions.

Proof.

$$\prod_{i=1}^{D}(2\pi\sigma_i)^{\frac{1}{2}}exp\{-\sum_{i=1}^{D}\frac{1}{2\sigma_i^2}(x_i-\mu_i)^2\}$$

□

Problem 3.4 (b). Show that the components of a random vector following a diagonal Gaussian distribution are independent random variables.

Proof. First, let us walk through the idea. We would like to prove the independence by proving that the product of each normal distribution's density function made up of the each component of the random vector is equal to the diagonal Gaussian distribution's density function, namely the random vector's density distribution.

The Gaussian density function for dataset D is defined as

$$p(x) = \mathcal{N}_{\mu,\Sigma}(x) = \frac{1}{(2\pi)^{d/2}\sqrt{det(\Sigma)}}exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu))$$

Since we have a diagonal Σ, it can be written in the form of a summation in the power of exponent, and then transformed into a product of exponents

$$p(x) = \frac{1}{(2\pi)^{d/2}\sqrt{det(\Sigma)}}$$

□