MLE estimates for a Gaussian distribution

1/1 point (graded)

In this problem, we will derive the maximum likelihood estimates for a Gaussian model.

Let X be a Gaussian random variable in d-dimensional real space (R^d) with mean μ and standard deviation σ.

Note that μ, σ are the parameters of a Gaussian generative model.

Recall from the lecture that, the probability density function for a Gaussian random variable is given as follows:

$$f_X\left(x|\mu,\sigma^2\right) = \frac{1}{\left(2\pi\sigma^2\right)^{d/2}} e^{-\|x-\mu\|^2/2\sigma^2}$$

Let $S_n = \{X^{(1)}, X^{(2)}, \ldots X^{(t)}\}$ be i.i.d. random variables following a Gaussian distribution with mean μ and variance σ^2.

Then their joint probability density function is given by

$$\Pi_{t=1}^{n} P\left(x^{(t)}|\mu,\sigma^2\right) = \Pi_{t=1}^{n} \frac{1}{\left(2\pi\sigma^2\right)^{d/2}} e^{-\|x^{(t)}-\mu\|^2/2\sigma^2}$$

Taking logarithm of the above function, we get

$$\log\{\Pi_{t=1}^{n}\frac{1}{(2\pi\sigma^2)^{d/2}}e^{-\|x^{(t)}-\mu\|^2/2\sigma^2}\}$$

$$=\sum_{t=1}^{n}\log\frac{1}{(2\pi\sigma^2)^{d/2}}+\sum_{t=1}^{n}\log e^{-\|x^{(t)}-\mu\|^2/2\sigma^2}$$

$$=\sum_{t=1}^{n}-\frac{d}{2}\log(2\pi\sigma^2)+\sum_{t=1}^{n}\log e^{-\|x^{(t)}-\mu\|^2/2\sigma^2}$$

$$\log P(S_n|\mu,\sigma^2)=-\frac{nd}{2}\log(2\pi\sigma^2)-\frac{1}{2\sigma^2}\sum_{t=1}^{n}\|x^{(t)}-\mu\|^2$$

Compute the partial derivative $\frac{\partial\log P(S_n|\mu,\sigma^2)}{\partial\mu}$ using the above derived expression for $P(S_n|\mu,\sigma^2)$.

Choose the correct expression from options below.

- $\frac{\partial\log P(S_n|\mu,\sigma^2)}{\partial\mu}=-\frac{1}{\sigma^2}\sum_{t=1}^{n}(x^{(t)}-\mu)$
- $\frac{\partial\log P(S_n|\mu,\sigma^2)}{\partial\mu}=\frac{1}{\sigma^2}\sum_{t=1}^{n}(x^{(t)}-\mu)$ ✔
- $\frac{\partial\log P(S_n|\mu,\sigma^2)}{\partial\mu}=\frac{1}{\mu^2}\sum_{t=1}^{n}(x^{(t)}-\mu)$
- $\frac{\partial\log P(S_n|\mu,\sigma^2)}{\partial\mu}=-\frac{1}{\mu^2}\sum_{t=1}^{n}(x^{(t)}-\mu)$

Solution:

$$\frac{\partial}{\partial\mu}\log P(S_n|\mu,\sigma^2)$$

$$=-\frac{1}{2\sigma^2}\sum_{t=1}^{n}-2(x^{(t)}-\mu)$$

$$=\frac{1}{\sigma^2}\sum_{t=1}^{n}(x^{(t)}-\mu)$$

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

MLE for the mean

1/1 point (graded)

Use the answer from the previous problem in order to solve the following equation

$$\frac{\partial \log P(S_n|\mu,\sigma^2)}{\partial \mu} = 0$$

Compute expression for $\hat{\mu}$ that is a solution for the above equation.

Choose the correct expression from options below

- $\hat{\mu} = \Pi_{t=1}^{n} x^{(t)}$
- $\hat{\mu} = \frac{\Pi_{t=1}^{n} x^{(t)}}{n}$
- $\hat{\mu} = \sum_{t=1}^{n} x^{(t)}$
- $\hat{\mu} = \frac{\sum_{t=1}^{n} x^{(t)}}{n}$ ✔

Solution:

Recall from the previous solution that

$$\frac{\partial \log P(S_n|\mu,\sigma^2)}{\partial \mu} = \frac{1}{\sigma^2}\sum_{t=1}^{n} \|x^{(t)} - \mu\|$$

Setting the above expression to zero, we get:

$$\frac{1}{\sigma^2}\sum_{t=1}^{n} \|x^{(t)} - \hat{\mu}\| = 0$$

$$\sum_{t=1}^{n} \|x^{(t)} - \hat{\mu}\| = 0$$

$$\sum_{t=1}^{n} (x^{(t)}) - n\hat{\mu} = 0$$

Resulting in the final expression for $\hat{\mu}$ as follows:

$$\hat{\mu} = \frac{\sum_{t=1}^{n} x^{(t)}}{n}$$

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

MLE for the variance

1/1 point (graded)

Compute the partial derivative $\frac{\partial \log P(S_n|\mu,\sigma^2)}{\partial \sigma^2}$ using the above derived expression for $P(S_n|\mu,\sigma^2)$ which is restated below as well:

$$\log P\left(S_n|\mu,\sigma^2\right) = -\frac{nd}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\left\|x^{(t)}-\mu\right\|^2$$

Choose the correct expression from options below.

- $\frac{\partial \log P(S_n|\mu,\sigma^2)}{\partial\sigma^2} = \frac{nd}{2\sigma^2} + \frac{\sum_{t=1}^{n}\|x^{(t)}-\mu\|^2}{2(\sigma^2)^2}$
- $\frac{\partial \log P(S_n|\mu,\sigma^2)}{\partial\sigma^2} = -\frac{nd}{2\sigma^2} + \frac{\sum_{t=1}^{n}\|x^{(t)}-\mu\|^2}{2(\sigma^2)^2}$ ✔
- $\frac{\partial \log P(S_n|\mu,\sigma^2)}{\partial\sigma^2} = \frac{nd}{2\sigma^2} - \frac{\sum_{t=1}^{n}\|x^{(t)}-\mu\|^2}{2(\sigma^2)^2}$
- $\frac{\partial \log P(S_n|\mu,\sigma^2)}{\partial\sigma^2} = \frac{\sum_{t=1}^{n}\|x^{(t)}-\mu\|^2}{2(\sigma^2)^2}$

Solution:

$$\frac{\partial \log P\left(S_n|\mu,\sigma^2\right)}{\partial\sigma^2} = \frac{\partial}{\partial\sigma^2}\{-\frac{nd}{2}\log\left(2\pi\sigma^2\right)\} - \frac{\partial}{\sigma^2}\{\frac{1}{2\sigma^2}\sum_{t=1}^{n}\left\|x^{(t)}-\mu\right\|^2\}$$

$$\frac{\partial \log P\left(S_n|\mu,\sigma^2\right)}{\partial\sigma^2} = -\frac{nd}{2\sigma^2} + \frac{\sum_{t=1}^{n}\left\|x^{(t)}-\mu\right\|^2}{2(\sigma^2)^2}$$

Submit

You have used 1 of 2 attempts

Answers are displayed within the problem

MLE for the variance

1/1 point (graded)

Using the answer from the previous problem in order to solve the equation

$$\frac{\partial \log P\left(S_n|\mu,\sigma^2\right)}{\partial\sigma^2} = 0$$

Compute expression for $\hat{\sigma}^2$ that is a solution for the above equation.

Choose the correct expression from options below

- $\hat{\sigma}^2 = \frac{\sum_{t=1}^{n}(x^{(t)}-\mu)^2}{nd}$ ✔
- $\hat{\sigma}^2 = -\frac{\sum_{t=1}^{n}(x^{(t)}-\mu)^2}{nd}$
- $\hat{\sigma}^2 = -\frac{\sum_{t=1}^{n}(x^{(t)}-\mu)^2}{n}$
- $\hat{\sigma}^2 = -\frac{\Pi_{t=1}^{n}(x^{(t)}-\mu)^2}{nd}$

Solution:

Recall from the previous solution that

$$\frac{\partial \log P(S_n|\mu,\sigma^2)}{\partial \sigma^2} = -\frac{nd}{2\sigma^2} + \frac{\sum_{t=1}^{n} \left\|x^{(t)} - \mu\right\|^2}{2(\sigma^2)^2}$$

Setting the above expression to zero, we get:

$$-\frac{nd}{2\sigma^2} + \frac{\sum_{t=1}^{n} \left\|x^{(t)} - \mu\right\|^2}{2(\sigma^2)^2} = 0$$

$$nd = \frac{\sum_{t=1}^{n} \left\|x^{(t)} - \mu\right\|^2}{\sigma^2}$$

The above equation leads us to our final expression for $\hat{\sigma}^2$:

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{n} \left(x^{(t)} - \mu\right)^2}{nd}$$

Submit

You have used 1 of 2 attempts

Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 4 Unsupervised Learning (2 weeks) :Lecture 15. Generative Models / 10. MLEs for Gaussian distribution