Audit Access Expires Dec 22, 2020
You lose all access to this course, including your progress, on Dec 22, 2020.

10. MLE for Gaussian Distribution

Bookmark this page

Exercises due Nov 17, 2020 18:59 EST Completed

MLEs for Gaussian Distribution

0:00 / 8:17 1.25x

Start of transcript. Skip to the end.

So now, imagine to yourself you are actually
given some training set.
And we want to find, as what we've
done before when we were talking about the question
of estimation, how can they find the best mu and sigma
squared that will give the highest
likelihood to my training data?
Again, in this case, the training data

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

MLE for the 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:

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

Let $S_n = \{X^{(1)}, X^{(2)}, \ldots X^{(n)}\}$ 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

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

Taking logarithm of the above function, we get

$$\begin{aligned} \log P(S_n|\mu,\sigma^2) = \log\left(\prod_{t=1}^{n} \frac{1}{(2\pi\sigma^2)^{d/2}} e^{-\|x^{(t)}-\mu\|^2/2\sigma^2}\right) &= \sum_{t=1}^{n} \log \frac{1}{(2\pi\sigma^2)^{d/2}} + \sum_{t=1}^{n} \log e^{-\|x^{(t)}-\mu\|^2/2\sigma^2} \\ &= \sum_{t=1}^{n} -\frac{d}{2}\log(2\pi\sigma^2) + \sum_{t=1}^{n} \log e^{-\|x^{(t)}-\mu\|^2/2\sigma^2} \\ &= -\frac{nd}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{n} \|x^{(t)}-\mu\|^2. \end{aligned}$$

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

Choose the correct expression from options below.

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

✔

Show answer

Submit You have used 2 of 2 attempts

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

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

✔

Save Show answer

Submit You have used 1 of 2 attempts

MLE for the Variance I

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 $\log P(S_n|\mu,\sigma^2)$ which is restated below as well:

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

Choose the correct expression from options below.

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

✔

Save Show answer

Submit You have used 1 of 2 attempts

MLE for the Variance II

1/1 point (graded)

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

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

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

Choose the correct expression from options below

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

✔

Save Show answer

Submit You have used 1 of 2 attempts

Discussion

Hide Discussion

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

Add a Post

Show all posts | by recent activity

General question (6)
Just one general thing. Anybody knows why we can't see the progress graph anymore? Thanks in advance.

Big picture question (3)
The start of this lecture the prof says "find the best u and sigma squared tha twill give highest likelihood to training data". What exactly does she mean b...

Squared norm of vector (2)
Hi folks Just wondering if anyone can explain how to get the derivative of the squared norm of a vector? What derivative rules do you need to apply? Tha...

MLE for Variance 1 (3)
Dear Staff, Kindly help in calculating second part derivate of 1st term It should be nd/(sigma^2) Because 1st derivative = -(nd/2)*(1/2sigma^2)*(4sigma)...