

<u>Unit 4 Unsupervised Learning (2</u>

Project 4: Collaborative Filtering via

Course > weeks)

> Gaussian Mixtures

> 5. Bayesian Information Criterion

## 5. Bayesian Information Criterion

So far we have simply set the number of mixture components K but this is also a parameter that we must estimate from data. How does the log-likelihood of the data vary as a function of K assuming we avoid locally optimal solutions?

To compensate, we need a selection criterion that penalizes the number of parameters used in the model. The Bayesian information criterion (BIC) is a criterion for model selection. It captures the tradeoff between the log-likelihood of the data, and the number of parameters that the model uses. The BIC of a model M is defined as:

$$\mathrm{BIC}\left(M\right) = l - \frac{1}{2}p\log n$$

where l is the log-likelihood of the data under the current model (highest log-likelihood we can achieve by adjusting the parameters in the model), p is the number of adjustable parameters, and n is the number of data points. This score rewards a larger log-likelihood, but penalizes the number of parameters used to train the model. In a situation where we wish to select models, we want a model with the highest BIC.

## Implementing the Bayesian Information Criterion

```
1.0/1.0 point (graded)
```

Fill in the missing Bayesian Information Criterion (BIC) calculation (bic function) in common.py.

**Available Functions:** You have access to the NumPy python library as <code>np</code>, to the <code>GaussianMixture</code> class and to typing annotation <code>typing.Tuple</code> as <code>Tuple</code>.

```
1 def bic(X: np.ndarray, mixture: GaussianMixture,
          log likelihood: float) -> float:
 2
      """Computes the Bayesian Information Criterion for a
 3
      mixture of gaussians
 5
 6
      Args:
7
          X: (n, d) array holding the data
 8
          mixture: a mixture of spherical gaussian
 9
          log likelihood: the log-likelihood of the data
10
11
      Returns:
12
          float: the BIC for this mixture
13
14
      n, d = X.shape
      K, = mixture.mu.shape
     return log likelihood - ((2+d)*K-1)*nn log(n)/2
```

Press ESC then TAB or click outside of the code editor to exit

Correct

### Test results

<u>Hide output</u>

CORRECT

est: bic fixed			
tput:			

```
Input:
X: [[0.85794562 0.84725174]
 [0.6235637 0.38438171]
 [0.29753461 0.05671298]
 [0.27265629 0.47766512]
 [0.81216873 0.47997717]
 [0.3927848 0.83607876]
 [0.33739616 0.64817187]
 [0.36824154 0.95715516]
 [0.14035078 0.87008726]
 [0.47360805 0.80091075]
 [0.52047748 0.67887953]
 [0.72063265 0.58201979]
 [0.53737323 0.75861562]
 [0.10590761 0.47360042]
 [0.18633234 0.73691818]]
K: 6
Mu: [[0.6235637 0.38438171]
 [0.3927848 0.83607876]
 [0.81216873 0.47997717]
 [0.14035078 0.87008726]
 [0.36824154 0.95715516]
 [0.10590761 0.47360042]]
Var: [0.10038354 0.07227467 0.13240693 0.12411825 0.10497521 0.12220856]
P: [0.1680912  0.15835331  0.21384187  0.14223565  0.14295074  0.17452722]
LL:-1067.804029
Output:
bix:-1098.946606
```

Test: bic random

#### Output:

```
Input:
X: [[0.32576222 0.6489857 ]
 [0.87380538 0.25631963]
 [0.93933927 0.61180819]
 [0.61864421 0.85014419]
 [0.67027679 0.76351659]
 [0.31981245 0.93619342]
 [0.93757478 0.05859924]
 [0.27851414 0.82655346]
 [0.95084238 0.67288828]
 [0.40109464 0.94913272]
 [0.06444047 0.60397294]
 [0.33430927 0.81271876]]
K: 5
Mu: [[ 0.33742528  0.31184246]
 [ 0.17439234 -0.07102599]
 [ 0.48474429  0.25333777]
 [ 0.62677286 -0.52945021]
 [ 0.45739841 -0.83798554]]
Var: [0.16462205 0.42297571 0.16517815 0.79397126 1.21333233]
P: [0.19528541 0.23274431 0.1611051 0.21943356 0.19143162]
LL:-1673.419842
Output:
bix:-1697.026455
```

<u>Hide output</u>

Submit

You have used 3 of 20 attempts

# Picking the best K

1/1 point (graded)

Find the best K from [1,2,3,4] on the toy dataset. This will be the K that produces the optimal BIC score. Report the best K and the corresponding BIC score. Does the criterion select the correct number of clusters for the toy data?

Best 
$$K = \boxed{3}$$

Submit

You have used 2 of 10 attempts

✓ Correct (1/1 point)

### Discussion

**Topic:** Unit 4 Unsupervised Learning (2 weeks): Project 4: Collaborative Filtering via Gaussian Mixtures / 5. Bayesian Information Criterion

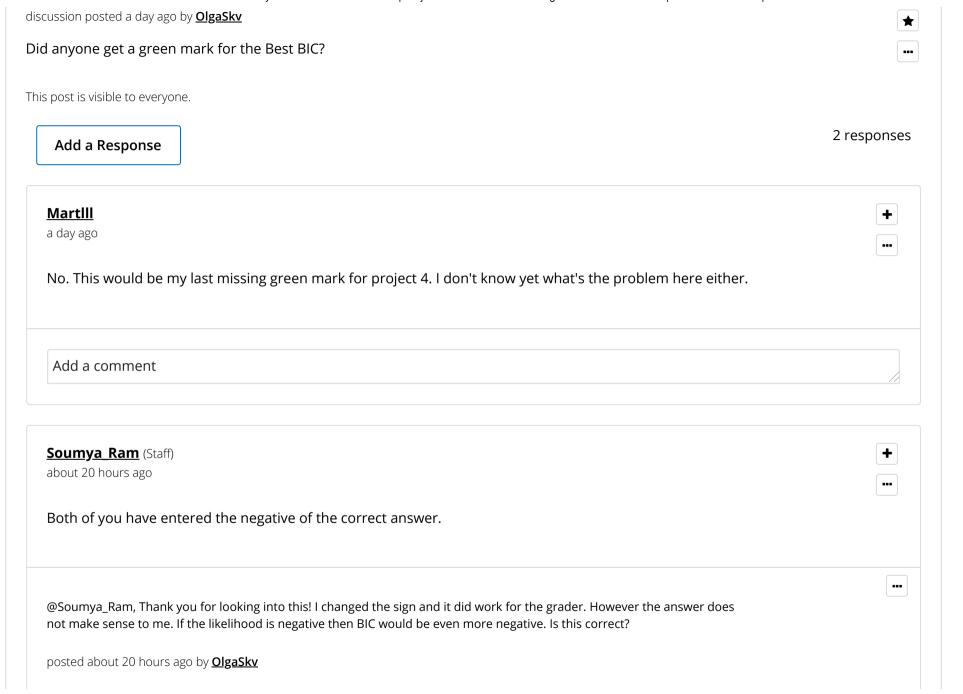
**Hide Discussion** 

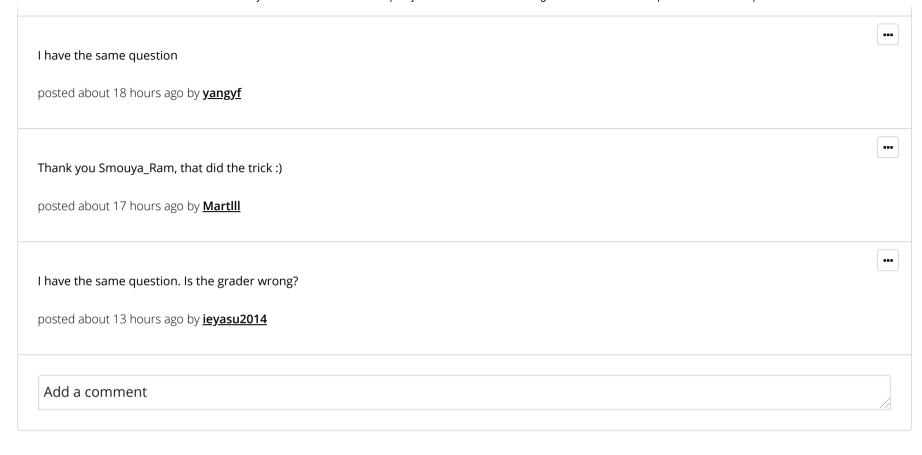
Add a Post

**≮** All Posts

**Best BIC** 







Showing all responses

Add a response:

8/6/2019	5. Bayesian Information Criterion   Project 4: Collaborative Filtering via Gaussian Mixtures   6.86x Courseware   edX
	Preview
	Submit

© All Rights Reserved