**GMDL, HW#2**אורן יקואל – 208727164 , יונתן לויטס - 336132014, גל אלון – 318512910

**Problem 1:**

We’ll show that s.t .   
As we saw in class we know that:  
   
Let’s denote a LaGrange multiplier function, with the constraints:  
So we’ll partially derive:  
 and   
So:   
   
And:   
together:  
   
We’ll recall,

**Problem 2:  
(i)** We’ll notice that if we choose we’ll get:  
Because we get an expression which doesn’t involve , and is a close expression, we can say is sampled from a uniform distribution.

**(ii)**  Uniform On the other hand , that is sampled from uniform distribution is (as proven above):   
Meaning, a uniform is a vector where entry is and , while a that is sampled from uniform distribution is a vector that can receive any value with equal probability.

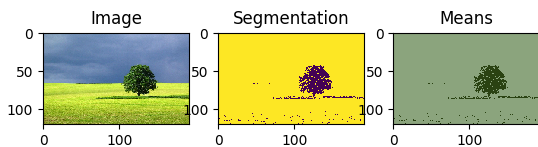
**(iii)** In the context of mixture models and estimating π , using the prior instead of a pure likelihood-based approach can prevent overfitting or misleading results. In particular, if the data set is too small or noisy, the prior can provide a useful source of information about the distribution, which can help better generalize new data.

For example, if we sampled 𝑁 (a relatively small number) samples from a mixture model and received data in which no sample comes from the entry with the largest weight – we’ll denote as c, then a likelihood-based calculation will give . On the other hand, if we can assume in advance that the data is sampled from some distribution where gets a relatively high value with a high probability, then the model will give a more accurate value for .

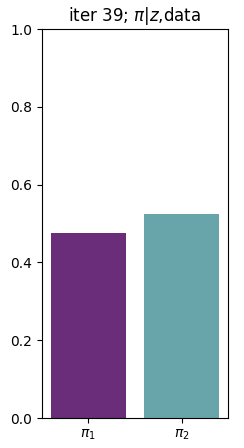
**Part 2:**  
Obviously, and as we saw the K parameter affects on how many different objects the model is trying to segment.

A picture containing text, screenshot, diagram, number

Description automatically generated**K = 2:**



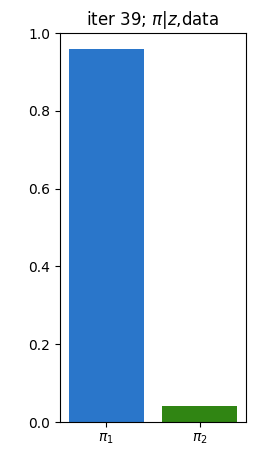
kappa = 1  
nu = 1000  
alphas = 100  
psi = 0.2  
m = 0.5

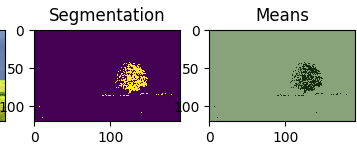


A picture containing text, screenshot, diagram

Description automatically generated

kappa = 1  
nu = 100  
alphas = 100  
psi = 0.2  
m = 0.5

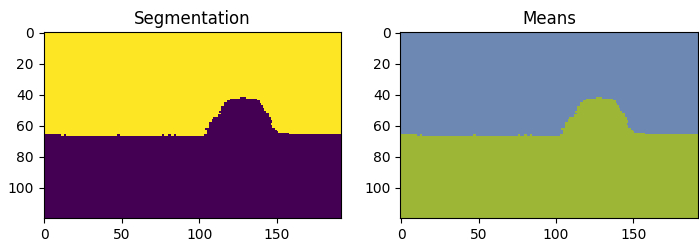




kappa = 1  
nu = 100  
alphas = 1000  
psi = 0.38  
m = 0.5

A picture containing text, screenshot, diagram, rectangle

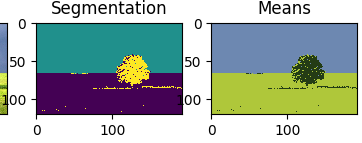
Description automatically generated



kappa = 1  
nu = 1000  
alphas = 100  
psi = 0.2  
m = 0.5

A picture containing text, screenshot, diagram, plot

Description automatically generated**K = 3:**



kappa = 10  
nu = 10000  
alphas = 100  
psi = 0.1  
m = 0.5

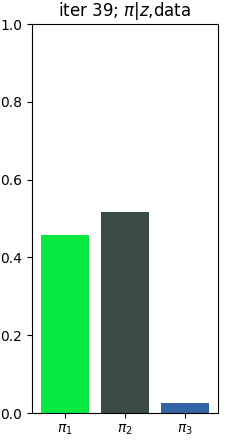
A picture containing text, screenshot, diagram, rectangle

Description automatically generated

A picture containing text, screenshot, tree

Description automatically generated

kappa = 1  
nu = 10000  
alphas = 100  
psi = 0.1  
m = 0.5



A picture containing text, screenshot

Description automatically generated

kappa = 1  
nu = 10000  
alphas = 100  
psi = 0.1  
m = 0.5

A picture containing text, screenshot, diagram, plot

Description automatically generated

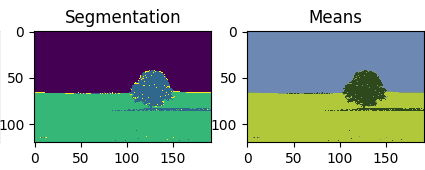
A picture containing text, tree, screenshot

Description automatically generated

kappa = 1  
nu = 10000  
alphas = 10000  
psi = 0.1  
m = 0.5

A picture containing text, screenshot, diagram, plot

Description automatically generated**K = 4:**



kappa = 1  
nu = 100  
alphas = 1000  
psi = 0.2  
m = 0.5

A picture containing text, screenshot, diagram, line

Description automatically generated

A picture containing text, screenshot

Description automatically generated

kappa = 1  
nu = 1000  
alphas = 1000  
psi = 0.3  
m = 0.5

A picture containing text, screenshot, diagram, plot

Description automatically generated

A picture containing text, screenshot, tree

Description automatically generated

kappa = 1  
nu = 5  
alphas = 1000  
psi = 0.2  
m = 0.5

A picture containing text, screenshot, diagram, plot

Description automatically generated

kappa = 1000  
nu = 100  
alphas = 1000  
psi = 0.2  
m = 0.5

A picture containing text, screenshot, christmas tree

Description automatically generated

**Analysis of hyper parameters**

Alphas - Dirichlet Hyperparameters, higher values of alphas result in a more peaked prior distribution, favouring fewer dominant components in the image segmentation. A higher alpha can influence the likelihood of gaussian centre to be selected and create segments bigger than they should be.

Kappa – Controls the strength of the prior belief, larger kappa is more data-driven when smaller kappa is more faithful to the prior. We can see that for a larger kappa it shifts the weight towards the larger shapes in the image.

Nu – Represents the degree of freedom in the hyperparameters, seems that a higher value of nu allows more flexibility in the segmented shapes.

Psi - Represents the prior precision matrix of the covariance matrices, a lower psi value results in more detailed segmentation.

M – Represents the prior mean vector – couldn’t find any visible affect it has.