Computational Cognitive Science (2017–2018)

School of Informatics, University of Edinburgh
Original exercises by Frank Keller, with modifications by Chris Lucas

Tutorial 2: Maximum Likelihood Estimation

Please work through this tutorial sheet on your own time as much as possible before arriving in tutorial. We encourage you to work together and discuss your methods and solutions.

1 Maximum Likelihood Concepts

Imagine you have conducted an experiment on memory recall. You have developed a novel model that you believe can explain the behavior you found in the experiment. Instances of your model are specified by a set of parameters.

Question 1: Within this context, what is the goal of maximum likelihood estimation (MLE)?

Question 2: Briefly describe an alternative way of fitting model parameters and compare it to MLE. Why might you choose one over the other?

Question 3: Could you compute the probability of a particular set of parameter values for your model, given the data, using only its likelihood function? If so, how? If not, why not?

2 MLE on the Ex-Gaussian Function

Download the following file, which contains the model that you will work with in this tutorial (it is linked from the Assignments page):

http://www.inf.ed.ac.uk/teaching/courses/ccs/tutorials/ccs_t02.tar.gz

Unpack this file, which will result in three Matlab source files:

- `exGaussPDF.m` implements the ex-Gaussian function introduced in the lecture;
- `Lopt.m` defines the joint likelihood function for the ex-Gaussian function and uses the Nelder-Mead algorithm to find its maximum;
- `LSurfaceL.m` plots the likelihood surface and runs the maximization.

The code is taken from chapter 4 of L&F and explained there in more detail. Please read this chapter if you need help understanding the code.

Now run `LSurfaceL.m`, which will generate a plot of the likelihood surface. You can examine it using the "Rotate 3D" button in the Matlab GUI.

Question 4: What is the shape of the likelihood surface? Try to estimate the maximum by eyeballing the plot.

Running `LSurfaceL.m` also runs the Nelder-Mead optimization algorithm and outputs both the operations performed and the optimum the algorithm finds.

Question 5: Examine the sequence of operations. Does the algorithm generate plausible output? Does it converge on a minimum for the likelihood function? Is the minimum the same as the one you estimated by eyeballing the plot?

Look at the code of `LSurfaceL.m`. In line 30, it calls the function `Lopt` which runs the maximum likelihood optimization. The start values for the parameters are $(\mu, \tau) = [0.1\ 0.1]$.

Question 6: Vary the start values of the parameters. Does the algorithm always converge on the same answer? Which results do you get for [0 1], [1 0], [1 2], [0 0]?

3 The Role of the Data for MLE

We will now investigate how the likelihood function $L(\theta|y)$ is influenced by the data y used to compute it. The code for `LSurfaceL.m` specifies `rt`, the vector of data points on which the likelihood estimation is carried out.

Question 7: What happens if the data vector `rt` is changed to contain a single point? Try `rt = [0.1]`, `rt = [1]`, `rt = [10]`. Explain the change in the likelihood surface. What happens for `rt = [100]`?

As explained in the lecture, in the general case, we will have multiple data points, and we are optimizing the joint likelihood of all the data points.

Question 8: Plot the likelihood surfaces for the following data vectors and compare them the ones you obtain if you have only a single data points (as in the previous question).

```
rt = [1 1 1 1 1 1 1 1 1 1 1 1]
rt = [1 1 1 1 1 1 1 1 1 1 1 10]
rt = [1 1 1 1 1 1 10 10 10 10 10 10]
rt = [3 4 5 6 6 7 7 8 8 9 10 11]
```

In the current implementation, the maximum likelihood estimation for the ex-Gaussian function only adjusts two parameters, μ and τ. Extend the code for `Lopt.m` and `LSurfaceL.m` so that the third parameter σ is also optimized.

Question 9: Run the new version of the code on the data vectors from the previous questions. Do you obtain a better fit using three parameters? How can you tell if your fit has improved?

Question 10: Can you think of a way of plotting a likelihood surface that depends on three parameters?

References

Lewandowsky, S. & Farrell, S. (2011). *Computational modeling in cognition: Principles and practice*. Thousand Oaks, CA: Sage.