Mathematical Methods for Neural Science and Engineering

Coursework Three
Deadline: 1pm, Mon 9th Jan, 2022.

Each coursework consists of a set of "real" data analysis or mathematical modelling tasks, which you should be able to carry out with the aid of what you have learned from the lectures and problem classes.

Your coursework should be created as a Jupyter Notebook. To submit this through Blackboard, first convert it to html. To do this, from the notebook itself run the following command:

```
! jupyter nbconvert --to html your_notebook_name.ipynb
```

Your coursework submission should include markup text describing your answers to the questions, and explanations, your code, and inline figures. Your figures should be of publication quality – label all axes, provide legends where appropriate, etc. Each answer should include a brief discussion of you're the meaning of your quantitative results, markup text.

You may collaborate, but you must write up your report **completely independently** (this will be checked), and reference all sources beyond the material on Blackboard.

Coursework Three

1. Load the calcium imaging dataset provided in the directory dataset3_m62 using the skeleton code provided ('skel_coursework3.py'), as in Coursework 1 and 2.
    a. You may have noticed that the calcium fluorescence time series for each cell in this dataset comprises brief "calcium transient" events interspersed amongst the baseline fluorescence. Write a function to convert each time series into an event train, in which an (instantaneous) event corresponds to the beginning of a detected calcium transient. Your detection code might be as simple as threshold crossing, or it might be more complex. This event train should be a binary process (i.e. "1" where there is an event, "0" otherwise). Given that the calcium transients have a rapid rise followed by a slow decay, you should have one event per calcium transient at the rise time, rather than the event train remaining at "1" for the whole calcium transient pulse (as the information is in the *onset* of the calcium transient).
    b. For the first three cells, plot the continuous fluorescence time series as a function of time, with the times of detected events marked by a tick over or under-laid on the same plot.

[30 marks]

2. As described in Coursework One, the data is collected from a mouse running in repeated loops around a circular track.
    a. Write a function which returns the data for a cell in the form response[trial,position], where trial indicates which loop around the track the mouse is on, position indicates the binned spatial angle of the mouse's location, and response is 0 or 1 as above.
    b. Plot this as a `rastergram', with ticks indicating events on axes showing trials (y axis) vs angular location (x axis)) for three example cells showing high, medium and low

selectivity of the firing of the cell to the mouse's location on the circular track. Show this rastergram for the same three cells (cells 0,1,2).

[20 marks]

3. For these 3 cells, calculate the mutual information between the binary response variable *r* and spatial angle, *s*, where *s* is the spatial angle binned into 20 bins.
    a. Use the "plugin" estimator for mutual information (i.e. direct implementation using the Shannon formula), and write your own function for estimating mutual information. Hint: while there are several ways to do this, one way is to collect a list of the response (0,1) and spatial location (0 .. 19) values for all time bins, and then use that to calculate the empirical probability distribution P(*r*|*s*).
    b. Does this correspond to what you expect based on the previous plots?

[30 marks]

4. Comment on whether your estimates of mutual information above are likely to be biased, and if so, by how much?

[10 marks]

5. Calculate the spatial information rates for all the cells and plot the distribution. The x-axis of your distribution plot should be in units of c (take the information rate as the Shannon mutual information divided by the bin width in seconds).
    a. Show the rastergram for the cell with the highest spatial mutual information (from all 138 cells).
    b. What is it about the response of this cell which improves its information rate compared to the three examples you looked at in 3?

[10 marks]

Simon Schultz
13 December 2022