Course Module

Analysis of extracellular activity across layers during a perceptual task

Synopsis

A central goal in neuroscience is to relate the activity of neurons to a subject's behavior. In this module we will use approaches based on signal detection theory and apply them to population recordings across cortical layers from macaque visual cortex while the animals are performing a visual task. Additionally, we will introduce an analysis technique that allows for the identification of cortical layers in these recordings.

Supplemental Material

With this course module we provide an Introductory Lecture *Analysis of extracellular activity in sensory cortex across layers during a perceptual task* and additional data files for practical analysis.

The supplemental material for this teaching module is available online at the following URL: www.g-node.org/teaching/xxx

Requirements

Practical work on this course module requires that you run Matlab Version 7 or higher. Supplemental data files are required for data analysis.

Practical work

Most decisions involve a degree of uncertainty. The uncertainty often results from noise, e.g. whether these are perceptual decisions (Was there a flash of light?), detection of a radar signal (Did the speed of this car exceed 50km/h?), or a classification algorithm ("Does this datapoint belong to class A or B?). Signal detection theory provides us with tools to analyze such decision under uncertainty. The concepts come from analyzing radar signals, were introduced to psychology in the 1950s, applied to neural signals in the 1990s and are widely used in neuroscience and machine learning to quantify classification performance.
In the following exercises you will be introduced to the basic concepts and apply them to behavioral data and simultaneously recorded neural data from visual area V2/V3a of a macaque monkey performing a visual task. I.e. the monkey has to make visually guided decisions based on noisy visual signals.

1) Introduction to signal detection theory

Let's assume you want to know whether your friend can hear a very faint beep on your phone. Your could play the beep many times and ask her whether she hears it, and to make sure that she tries really hard you tell her she will get 10 cents every time she hears it. But what about if she's cheating and says "yes" every time to get your 10 cents?

To prevent her from cheating, you could introduce no signal trials ("catch" trials) on which you don't play the beep. Now assume you play the beep only on half of the trials. Now, if she says 'yes' all the time or 'no' all the time, she will only 50% correct, at chance level. If the beep is very weak, she will not be correct on all the trials. So your friend is faced with a decision under uncertainty. Signal detection theory (SDT) is helpful in this situation.

SDT assumes that your friend bases her decision on an internal response to the signal. If this internal response exceeds a certain criterion (*c*), she will say yes, otherwise she will say no. There is some variability in this internal response.

We will assume that the variability has a Gaussian distribution (SD = 1) around a certain mean. Let's assume in response to just noise, the distribution of the internal response is centered around 0, and in response to the beep, it is centered around 1.

On each trial there are now four possible outcomes, and SDT categorizes them as follows:

Signal	**Response** Yes	No
Present	"hit"	"miss"
Absent	"false alarm"	"correct reject"

Tasks:

1) What could be potential sources of this variability in this assumed internal response?

2) Plot the probability distributions of the internal response to the noise and the signal distributions.

3) *Exploring the role of the criterion:*
   a) Assume that the criterion is 1. Superimpose it as a vertical line in your plot.

   b) How are the probabilities of a Hit, Miss, False alarm (FA) and correct reject (CR) reflected in the plot?

   c) What are the probabilities of a Hit, Miss, False alarm (FA) and correct reject (CR) for this criterion?

   d) Assume that you give your friend 10cents for hits but she has to pay 50cents for every FA (Misses and CR are neither rewarded nor punished). What effect might this have on her criterion?
   Or think about a radiologist looking at an X-ray to check for lung cancer. What should the criterion be in this case?

   e) Play around with different criterion levels and re-compute the probabilities. For which criterion is your friend's performance best (maximal proportion correct)? Compare this to the probability distributions in point 2).

4) *The receiver operating characteristic (ROC) curve:*
   a) You may have noticed when playing around with the criterion that there is a

tradeoff between the probabilities for a hit (hit rate) and false alarms (FA rate). What is this tradeoff?

b) Plot for each criterion that you explored the hit-rate (y-axis) against the FA-rate (x-axis).
Now do this more systematically and compute the Hit and FA-rate for evenly spaced criterion values, e.g. criteria=[-4:0.1:6]. Again, plot the hit-rate against the FA-rate. The resulting curve now is an ROC curve that summarizes the performance given the distribution of the internal response for the signal and the noise for all criterion levels.

c) What does the ROC curve look like if the distributions for the signal and the noise are identical? What would be the proportion correct in this case?

Now explore the ROC curve for different values for the mean and SD of the distributions. Try also other distributions, e.g. Poisson distribution. (Make sure that the mean for the signal distribution is higher than that for the noise. Or explore what happens if that's not the case.)

d) You hopefully noticed that by decreasing overlap between the signal and noise distributions (e.g. increasing the difference in the means or decreasing the SD), the ROC curve bowed further away from the diagonal. It means that the signal and noise can be better discriminated.

The area under the ROC curve (aROC) quantifies this discrimination performance. Compute it for the different distributions (use the MATLAB function trapz() to compute the area). Typically aROC values range from 0.5 (chance performance) to 1 (perfect discrimination).

5) *The aROC and the performance in a two-alternative forced choice task.*
Assume that instead of presenting one stimulus on each trial, you present your friend with two consecutive presentation intervals and ask on each trial, whether the beep occurred in the first or second interval. (Note, that this is the design of a classical two-alternative forced choice task in psychophysics.)
Assume further, that your friend chooses the interval on each trial for which the internal response was higher.
Run a small simulation to model her responses. (signal mean = 1; noise mean =0; SD =1). Simulate at least 1000 trials. What is the percent correct in the simulations? Compare this value to the aROC for these noise and signal distributions of the internal response. What do you notice?

6) *Optional 1:* What happens when aROC values are <0.5? Think about the ROC curve in that case and the signal and noise distributions.

Optional 2: MATLAB has a function "perfcurve" that performs most of the computations that you have just coded up. They also give a few examples illustrating common uses of ROC curves. You might want to check out the examples.

2) Using signal detection theory to analyse neuronal data

The matlab file **ch19.mat** contains extracellular spikes recorded from one contact (ch19) of a multichannel electrode inserted into visual area V2 of a macaque monkey while this monkey performed a visual discrimination task. (The discrimination is a depth discrimination task. On each trial the answer is either "near" or "far"). The stimulus is parametrized from -1 to 1, with -1 a strong "near signal" stimulus, 0 a pure noise stimulus (think of the trials without a beep, or "catch" trials in the example above), and 1 a strong "far signal" stimulus.

The data is summarized in a matrix, where each row corresponds to one trial, and each column to the following:
1: Trial number
2: Stimulus (ranging from -1 to 1)
3: neuronal response, i.e. spike count on this trial (in spikes/trial)
4: choice of the monkey (0= “near”; 1=”far”)

Tasks:

1) Take a look at the data.
    a) Sort the neuronal responses by stimulus and plot their distributions as histograms for each stimulus.
    b) Plot the mean spike count (+/-SE) as a function of the stimulus. What do you notice?

2) *The neurometric function:* Use the concepts of signal detection theory that we introduced above to quantify the discrimination performance of this neuron.

    a) How well does the neuron discriminate a 0 and a 0.5 stimulus? *Hint:* treat the spike counts as the internal responses in the above example and, and the 0 stimulus as the catch trials above (i.e. those without any beep). Compute the aROC.

    b) Now compute an aROC for each stimulus compared against the 0 stimulus. What is the aROC for the 0 stimulus?
    Plot the aROC against stimulus strength (stimulus on the abscissa). This function (aROC vs stimulus) is commonly referred to as a neurometric function.

3) *The psychophysical function:* Now compute the monkey’s performance for each stimulus. *Hint:* quantify performance as proportion “near” choices. Plot the performance as a function of the stimulus (“psychophysical function”). Compare the psychophysical function with the neurometric function you generated in 2b. How does the behavioral discrimination performance compare to that of the neuron?

4) *Choice probabilities:* Above, we analysed the neuronal data and the behavioral data separately. But ultimately we want to link the two, and since both, the behavior and the neuronal data were measured simultaneously, we may be in a position to do so.

    a) We will assume (almost correctly) that all stimuli of the same type are identical. So, if there is some variability in response to the same stimulus that means that it is due to something happening in the brain, e.g. internal noise, and might already be reflected at the level of V2. Use SDT to determine how well the V2 neuron can predict the animal’s choices (independent of the stimulus, i.e. restrict your analysis to one type of stimulus—start with the 0 stimulus).
    The aROC for this analysis here is typically referred to as “choice probability” (CP).

    b) Do the same analysis for all stimulus types.

    c) *Optional:* Calculations of CPs are more reliable when the number of responses (number of near vs number of far choices in this case) are approximately balanced. Use bootstrapping to verify this by computing errorbars for the CPs for each stimulus type.

3) (optional) Testing for optimal linear read-out of the visual information for this task and (optional advanced) DATA CHALLENGE.

Typically, choice probabilities in early sensory areas have been slightly but very consistently above chance level. This is an interesting finding by itself (reading the animal's mind based on a single neuron in their visual cortex) . But the main reason why this observation has received substantial attention is the hope that it allows one to infer how the brain is reading out the information from sensory neurons. This turned out to be surprisingly difficult, mostly because of co-variability between neurons (remember the "noise correlations" from Day2?). But Haefner et al. (2013) has developed nice analytical tools to address this problem. Pitkow et al (2015) used these to devise a test for optimal linear read-out in the presence of ("information-limiting") noise correlations.

Tasks:

1) Convert the CPs you computed into "choice correlations" as defined by Pitkow et al. (their equation 9)
2) Use their equation 3 to test for optimal linear read-out. *Hint:* Consider θ in their equation as the psychophysical threshold of the monkey during the experiment, and θ_k the neurometric threshold. To find these fit a cumulative Gaussian to the psychophisical and neurometric functions you computed in Tasks 2-2 and 2-3. The SD of the fitted cumulative Gaussian is the psychophysical/neurometric threshold. Use fminsearch to find the best fitting curve (ideally, using maximum likelihood estimation but you can use your favorite approach).
3) Do the same analysis for all units (ch1.mat through ch24.mat).
4) Do the data support optimal linear read-out?

5) *DATA CHALLENGE (Advanced optional):* Explore additional classifiers (there are some out-of the box algorithms available in MATLAB) to predict the monkey's choices based on all simultaneously recorded units. How well can you do? (Make sure to cross-validate the performance). How well can you do when you take out the predictive effect of the response modulation due to the stimulus?

4) Current-source density analysis to infer the position of layer 4

The above recordings were done using a linear array with 24 contacts that was inserted approximately vertically to the layers of V2. The recordings therefore come from units in different layers but whose receptive fields are approximately overlapping. However, without further analysis we do not know which contact is located in which layer. The following analysis is a tool to give us an approximate idea of this. Rather than extracellular spiking activity we will rely on the local field potential (LFP, voltage recorded at each contact and filtered between 1-250Hz, sampled at 1kHz) for this analysis.

The matlab file **lfp.mat** contains lfp signals recorded from each channel, aligned on stimulus onset from -200 ms before to 200ms after stimulus onset.
lfp is a [c x t x n] matrix, where:

c: number of channels (24)
t: time samples (401 samples: 200ms before to 200ms after stimulus onset)
n: number of trials (227)

Tasks:

1) Explore the raw data. Plot the LFP aligned on the stimulus onset. Think about ways how to best visualize the data.

2) Plot the average LFP for each channel, aligned on stimulus onset, and sort it by channel number. Use a colorplot (e.g. imagesc) to visualize the data. You have now a spatio (the inter-channel spacing is 50mu)-temporal plot of the stimulus triggered LFP.

3) The current source density map (CSD map) can be computed as the 2nd spatial derivative of the average stimulus-triggered LFP signal across channels. Compute it.

4) You will notice that the CSD map is quite noisy. It is therefore helpful to filter the stimulus triggered LFP both temporally and spatially.
   a) Bandpass filter the average LFP (e.g. 8-100 Hz, using a 4-pole low-pass butterworth filter and a 2-pole high-pass butterworth filter; use the MATLAB function butter.m to specify your filter, and filter.m to filter the signal). You can also explore different filter settings.
   b) Apply a spatial lowpass filter to the temporally filtered. (e.g. use a 2D Gaussian filter of SD 1.1 inter-channel distance; use conv2.m to filter the data). Make sure to duplicate the data for the first and last channel before filtering to reduce edge effects.
   c) Apply the Vaknin method (Vaknin et al, 1988) to avoid "edge effects" in the CSD computation. Vaknin et al. argue that above layer 1 and below layer 6 (outside the gray matter) the LFP signal changes little such that it is an acceptable approximation to just extrapolate the LFP signal by duplicating it for the first and last channel (position in Vaknin's case). Do this on the filtered LFP signal then compute the CSD map just like in point 3.

5) Inspect the CSD map. The first main current sink has been shown to reflect the location of the dominant feed-forward input. At which channel(s) would you locate it? Define the center of the first current sink as 0 and plot your CP results from 2_4a/b relative to this 0 position. Is there a systematic change of CP across layers?

Acknowledgements:
The introduction to signal detection theory is largely based on Geoffrey Boynton's SDT tutorial for the Cold Spring Harbor Course in Computational Neuroscience.