Mutual information in metric spaces

Kristian Krastev

Supervised by Dr. Conor Houghton

BSc Final-year project

Department of Computer Science

University of BRISTOL

Project outline

An elegant new formula for mutial information in spaces without coordinates was proposed by Dr. Conor Houghton, motivated by the problem of estimating information in the space of spike trains, but applicable to any metric space.

Mutual Information tells us how much knowledge of one variable reduces uncertainty about another one. It is a powerful tool for quantifying relationships between random variables. Like any information quantity, in practice it requires very large data samples in order to be accurately estimated using the standard equation. This necessitates the use of alternative methods for estimating probability densities, based on which information can be approximated.

Metric spaces

These are sets with a distance or dissimilarity measure defined between any pair of their members. Volumes in them are hard to estimate as there might be no meaningful coordinates. Probability however always gives a measure.

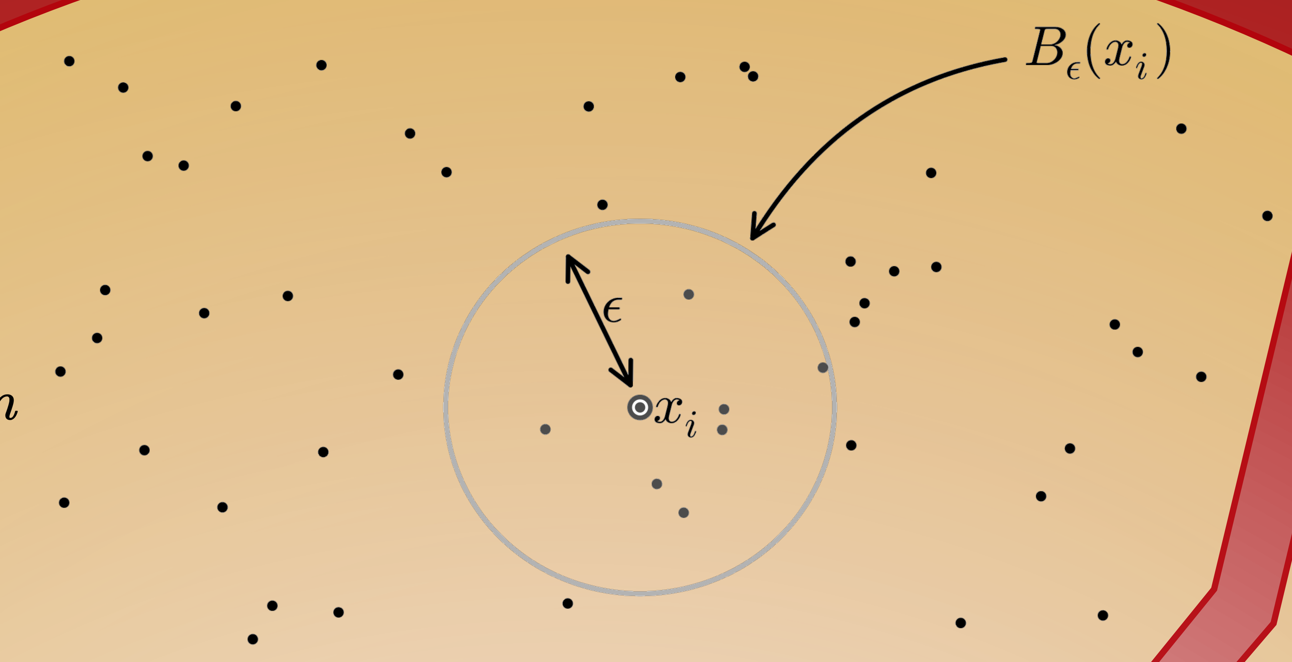
Neural signalling

Information is communicated between neurons as sequences of stereotypical events called action potentials or spikes. Both spike firing rates and exact spike timing convey transmitted information

Spike-train metrics

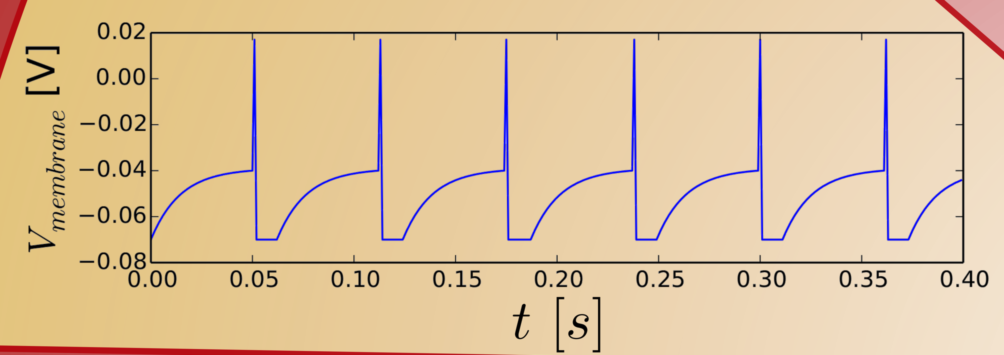
Breaking up spike trains into discrete time slots is inefficient because they occur at a time-scale as narrow as 1 ms, resulting in sparse high-dimensional spaces and excessive amounts of possible words over a meaningful period. There exist however other, more elaborate methods for embedding spike trains into a space. One approach uses edit-length distances, analogous to Hamming distance with a custom cost function, while another one maps spike trains to functions using a kernel, and subtracts them linearly to obtain a distance. Both induce non-coordinate metric spaces.

An open ball of some radius around a spike-train instance is defined in a metric space.



$$B_\epsilon(x_i) = \{t \in \mathcal{X} : d(x_i, t) < \epsilon\}$$

Spike trains are obtained by simulating membrane potentials.



$$V \approx \frac{\#[B]}{N}$$

Probability is used to measure volumes.

Assuming the PMF is constant in the ball, it is approximated as :

$$p_X(x_i) \approx \frac{\#[B(x_i, V)]}{NV}$$

ϵ is fixed s.t. B contains the h nearest neighbours of x_i, i.e. $V = \frac{h}{N}$

This measure cannot be used for entropy since probability is defined in terms of volume and vice versa:

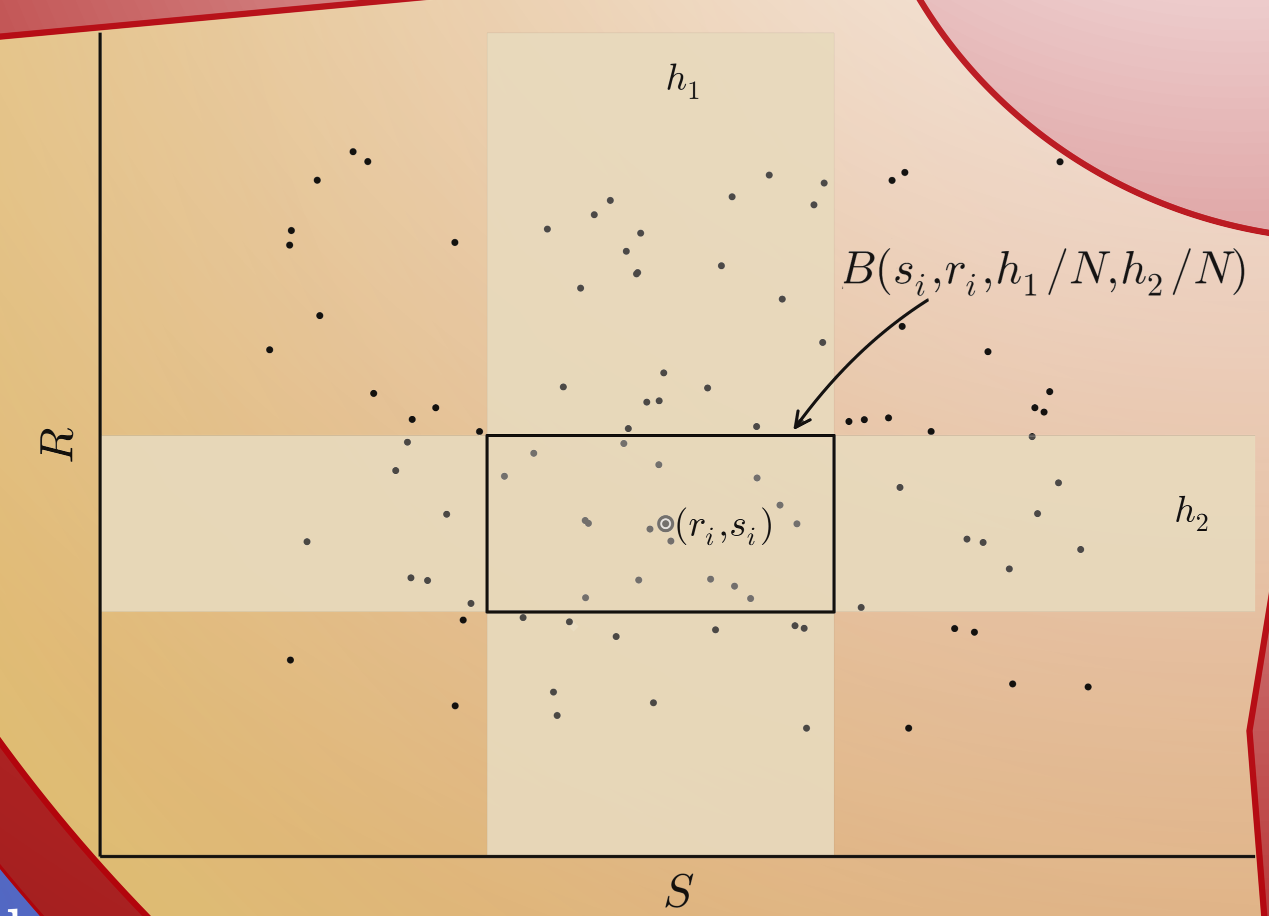
$$H(X) \approx -\frac{1}{N}\sum_{i=1}^{N} \log_2 p_X(x_i) = 0$$

But one probability distribution can be used as a measure to estimate others!

$$I(S;R) \approx \frac{1}{N}\sum_{i=1}^{N} log_2 \frac{N\#[B(s_i, r_i)]}{h_1 h_2}$$

Given N samples of two metric-space variables $S \in \mathcal{S}, R \in \mathcal{R}$:

$$B\left(s_i, r_i, \frac{h_1}{N}, \frac{h_2}{N}\right) = \left\{(s,r) \in \mathcal{S}\times\mathcal{R} : s \in B_S\left(s_i, \frac{h_1}{N}\right), r \in B_R\left(r_i, \frac{h_2}{N}\right)\right\}$$



The mutual information is estimated using the count of data points which fall in the cross-section of the h_1 nearest neighbours of s_i and the h_2 nearest neighbours of r_i as a probability estimate.

Current progress

A software simulating a network of integrate-and-fire neurons given a connectivity matrix has been implemented in Python. This produces outputs in the desired formats, which in turn serve as the basis for conducting experiments reflecting different network topologies.

The main amount of research on information estimation and neural modelling has been made.

Two types of metrics, which best reflect neural coding theories have been implemented.

Future work

The completion of the project consist of the follwing steps:

- A metric will be chosen in order to compute mutual information using the formula.
- The model will be applied to infere connectivity in a population of firing neurons using mutual information between their spike trains. This will test if it works.
- The resolution given by the size of the balls will be tuned to obtain optimal results.
- These results will then be compared with the ones obtained with other, more standard methods.

References

[1] Shannon CE. (1948) A mathematical theory of communication. *Bell Syst. Tech. J. 27: 379-423,623-656.*

[2] Kozachenko LF, Leonenko NN. (1987) Sample estimate of the entropy of a random vector. *Probl. Pered. Inf. 23: 9–16.*

[3] Kraskov A, Stogbauer H, Grassberger P. (2004) Estimating mutual information. *Phys. Rev. E 69, 066138.*

[4] Houghton CJ. (2015) Calculating mutual information for spike trains and other data with distances but no coordinates. *R. Soc. open sci. 2: 140391.*