Application of the Elbow Method to Determine Place Field Cluster Pattern along Dorsal-Ventral Hippocampal Axis

William Hockeimer

February 17, 2015

Motivation

Our understanding of how the mammalian brain solves the problem of spatial navigation has been greatly aided by the discovery of neurons within the medial temporal lobe with spatial receptive fields. The best studied are neurons located in the hippocampus, termed 'place cells', which fire at unique positions in the environment (O'Keefe 1976). The area in external space which reliably evokes firing of a given place cell is the place field of that cell. It is known that place field characteristics vary along the dorsal-ventral axis of the hippocampus such that the place fields become progressively larger going from dorsal to ventral positions (Jung et al., 1994; Kjelstrup et al., 2008). Thus there exist numerous maps of the current environment within the hippocampus, each with a different characteristic resolution. This multiplicity of scales may allow different sized environments to be encoded equally well, or it may allow more locations to be encoded by an 'address' system that proceeds from coarser maps to finer ones. From a graph theoretic perspective, an important question is how do these groups of place cells group themselves based on the size of their place field? Certainly they exist along a continuum but what degree, if any, of clustering occurs based on firing field size? More formally, what is the optimum number of clusters that describe place field organization based on firing field size along the dorsal-ventral hippocampal axis?

Overview of Method

As the central question is one of clustering, the k-means clustering algorithm will be used as a basic yet useful method of clustering observations. This algorithm requires the user to determine the number of clusters, however the number of clusters of place cells is in some sense the main question here, which is a bit of a Catch-22. To resolve this, increasing number of clusters will be used to explain simulated data with known labels and the elbow method will be used to determine when adding more clusters diminishes returns in explanatory power (see below for details).

Statistical Decision Theoretic

Each statistical decision problem has six features which describe it. The current question consists of applying a statistical decision problem a number of times while varying a key parameter (cluster number), so there is an additional step enumerated below the cardinal ones.

Sample Space $G_n = (V, E, Y)$

This describes the state space of all possible graphs where V = number of vertices = ~1E6, E = their connections = each has ~1E4 connections and Y = labels of location, generated pseudo-randomly based on empirical data of distribution of cells along D-V axis.

Model

$$SBM_n^k = (\rho, \beta)\; k = \alpha, \rho\epsilon\Delta_\alpha, \beta\epsilon(0,1)^{\alpha,\alpha}$$

The model is a stochastic block model broken up into alpha number of blocks (clusters) , each of which is associated with a different sized place field of the place cells grouped within that cluster. The number of blocks, k, is not constant but is varied with each iteration through the statistical decision problem. For example, the clustering algorithm is run on an SBM with one block (k=1) and the variance explained is computed. Then the clustering is done with k=2,k=3,…,k=x blocks until the optimal number of blocks to explain the clustering of the place cells is obtained. This k is a variable which is being optimized by some method, here the elbow method (below).

Action Space

$$A = \{y\,\epsilon\ \{0, ..., \alpha\}^n\}$$

The action space describes the possible outcomes of each decision rule instance, i.e. the possible ways an observation can be clustered. Here, each observation can be placed into one of alpha clusters.

Decision Rule Class

$$\underset{\mathbf{S}}{\arg\min} \sum_{i=1}^{k} \sum_{\mathbf{x}\in S_i} \left\|\mathbf{x} - \boldsymbol{\mu}_i\right\|^2$$

The decision rule is the clustering algorithm chosen to partition the observations into k number of clusters, where here k = alpha. The algorithm used here is the popular k-means clustering algorithm which attempts to minimize the within-cluster sum of squared error. Intuitively, it uses the cluster number k to establish k number of centroids and moves them about the state space until each observation is a minimal distance away from a centroid.

Loss Function

$$\sum_{i=1}^{n} \Theta(\hat{y}_i = y_i)$$

The loss function describes how many observations will be mis-categorized during each run of the statistical decision problem / clustering algorithm. The observations being clustered are generated from simulated data based on empirical results so the location of each place cell is known beforehand. Thus after the algorithm clusters each observation into some compartment along the D-V axis in the hippocampus, one can see how often that assignation matches the known label.

Risk Function

$$R = P \times l$$

The risk function is usually the expected loss of the function, i.e given the loss function above how much loss we actually expect.

Elbow Method

The elbow method runs a clustering algorithm multiple times with increasing numbers of clusters, i.e. on each run k = alpha where alpha is some integer that increases on each run through the algorithm (Thorndike, 1953). For each cluster number, the variance explained by using that number of clusters is computed. Thus one can plot variance explained vs. number of clusters. When this is done the graph will show a sharp increase for small numbers of clusters as large amounts of the variance start to be captured by the clusters. However, as the number of clusters increases there is are diminishing returns in the explanatory power of variance and at some point there is a pseudo-inflection point, an "elbow" that shows visually and intuitively when the amount of variance explained by further clustering is minimal. Thus the optimal number of clusters to use occurs at the elbow point. Note the elbow cannot always be identified (Ketchen and Shook, 1996).

Questions

1. Am I correct in understanding the blocks of the SBM as being potential clusters? Because I also remember you can have sub-graphs within them, I believe, so there could be further clustering within a block?
2. Would using the elbow method and increasing k on each run through the algorithm work? Fundamentally, have I confused some part of the process? And even if not, is it computationally tractable?

References

O'Keefe, John. "Place units in the hippocampus of the freely moving rat."*Experimental neurology* 51.1 (1976): 78- 109.

Kjelstrup, Kirsten Brun, et al. "Finite scale of spatial representation in the hippocampus." *Science* 321.5885 (2008): 140-143.

Jung, Min W., Sidney I. Wiener, and Bruce L. McNaughton. "Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat." *The Journal of neuroscience* 14.12 (1994): 7347-7356.

Robert L. Thorndike (December 1953). "Who Belongs in the Family?". *Psychometrika* 18 (4): 267–276

David J. Ketchen, Jr & Christopher L. Shook (1996). "The application of cluster analysis in Strategic Management Research: An analysis and critique". *Strategic Management Journal* 17 (6): 441–458.