Minimal embedding dimension bounds for receptive fields

Andrew Lampinen

May 7, 2014

Abstract

Curto, et. al., in 2013 outlined the neural ring, and neural ideal for describing receptive field structures of groups of neurons. In this paper we review their results, and explores some bounds that the algebraic and simplicial complex structures induce on the minimal embedding dimension of the receptive fields.

1 Introduction

In [1], the authors describe a way of representing the spatial structure of the stimulus regions that sensory neurons are attuned to, (receptive fields) in terms of algebraic properties. Specifically, they create a variety (neural code) from the subsets of the neurons all of which fire in response to given stimuli, and then construct the ideal of this variety explicitly, construct another related ideal, define a canonical representation for it, and give several algorithms for performing computations with it, including primary decompositions. One of their principle assumptions throughout the paper is the convexity of the receptive fields. This key assumption restricts which dimension of stimulus spaces can correspond to a given variety, as there are many set intersection and containment relations which cannot be realized in convex sets in a low dimensional space.

The authors construct a lower bound on the minimal embedding dimension, however they state that it is "crude," and it is constructed purely from the simplicial complex associated with the code, rather than exploiting the additional algebraic structure. In this paper we further explore these bounds, giving additional lower bounds as well as some illustrative counter-examples to possible alternative bounds.

2 The fruit of (receptive) fields

Although it would be impossible here to give a full exposition of the concepts of [1], in this section we will briefly review the products thereof which are essential to understanding this paper.

Receptive fields are maps $f_i : X \to \mathbb{R} \geq 0$, which map a stimulus space to the firing rate of a neuron. ([1]) Following the notation of [1], as we will throughout this paper, we denote by a receptive field both the map itself and its support, as the structure of the neural rings and ideals only depends on whether a neuron fires at all in response to a given stimulus.

Suppose we have a set of n receptive fields $\mathcal{U} = \{U_1, ..., U_n\}$ with the U_i subsets of some stimulus space X. We define the *Receptive Field code* (RF code) of $\mathcal{U}$, $C(\mathcal{U})$ to be the set of code words which are realized by some stimulus in X, i.e.

$$C(\mathcal{U}) = \{\bar{c} \in \{0,1\}^n : \exists x \in X \text{ with } x \in U_i \iff \bar{c}_i = 1\}$$

Fron $C(\mathcal{U})$, we may construct an abstract simplicial complex (if a set is in it, so are the set's subsets,)

$$\Delta(C) = \{\sigma \subset \{1, ..., n\} \mid \exists c \in C \; c_i = 1 \text{ if } i \in \sigma\}$$

Note that the definition involves is an if, not an iff. This is to ensure we actually produce a simplicial complex. We will sometimes refer to the simplicial complex $\Delta(C(\mathcal{U}))$ as the *nerve* of $\mathcal{U}$, $N(\mathcal{U})$. This is the set of all subsets of U which have non-empty intersection.

From a neural code $C(\mathcal{U})$, we may also define the ideal I_C by letting it be the ideal of polynomials which vanish on $C(\mathcal{U})$,

$$I_C = I(C(\mathcal{U})) = \{f \in \mathbb{F}_2[x_1, ..., x_n] : \forall c \in C, \; f(c) = 0\}$$

We define the neural ring R_C to be the coordinate ring of C, in the standard algebraic geometry sense, i.e.

$$R_C = \mathbb{F}_2[x_1, ..., x_n]/I_C$$

The ideal I_C contains the boolean relations $x_i(1-x_i)$ no matter what C is, because the field has characteristic 2, we would like to create a slightly smaller ideal which only encodes the relevant structure. We do this by defining for any $v \in \{0,1\}^n$ the characteristic function

$$\rho_v = \left(\prod_{\{i|v_i=1\}} x_i \right) \left(\prod_{\{j|v_j=0\}} (1-x_j) \right)$$

This function is 1 when evaluated at v, and 0 everywhere else. Then we define the neural ideal to be

$$J_C = \langle \rho_v : v \notin C(\mathcal{U}) \rangle$$

Then $J_C \subset I_C$. In [1] it is proven that in fact

$$I_C = J_C + \langle x_i(1-x_i) : i \in \{1, ..., n\} \rangle .$$

A *pseudo-monomial* is a function $f \in \mathbb{F}_2[x_1, ..., x_n]$ which is of the form $f = \prod x_i \prod(1-x_j)$ where no x_i appears in both products. (We define the degree for these to be the polynomial degree of f, this is in keeping with the standard definition of degree if f is a monomial.) In particular, it is clear that ρ_v are pseudo-monomials. An ideal $I \subset \mathbb{F}_2[x_1, ..., x_n]$ is a *pseudo-monomial ideal* if it can be generated by pseudo-monomials. Hence J is a pseudo-monomial ideal.

Finally, we define the *canonical form* of J_C, $CF(J_C)$ to be the set of all pseudo-monomials of J which are not multiples of a pseudo-monomial of lower degree. Then $J = \langle CF(J_C) \rangle$, and $CF(J_C)$ is unique. It is from this canonical form that we will attempt to create bounds on the embedding dimension of the receptive fields.

The authors of [1] partition the monomials which occur in $CF(J_C)$ into three broad categories (letting $\sigma, \tau \subset \{1, ..., n\}$ with $\sigma \cap \tau = \emptyset$):

1. Type I: Monomials of the form $\prod_{i\in\sigma} x_i$. This monomial's appearance implies that $\cap_{i\in\sigma} U_i = \emptyset$, but for any proper subset of σ the intersection is non-empty.

2. Type II: Monomials of the form $\prod_{i\in\sigma} x_i \prod_{j\in\tau}(1 - x_j$. These correspond to

$$\emptyset \neq \bigcap_{i\in\sigma} U_i \subseteq \bigcup_{j\in\tau} U_j$$

   and there exists no pair of subsets of σ and τ, at least one of which is proper, which satisfies this relationship.

3. Type III: Monomials of the form $\prod_{j\in\tau}(1 - x_j)$. These correspond to

$$\bigcup_{j\in\tau} U_j = X$$

   i.e. the $\{U_j | j \in \tau\}$ cover the stimulus space X.

3 Helly's theorem and other background

Before we continue, we will need some classical results. In this section we will state Helly's theorem on intersections of convex sets, a lemma, and a result by Kalai, all of which will be useful later. For the proof of Helly's theorem, see [2].

Helly's Theorem: Suppose $\mathcal{K} = \{K_1, ..., K_n\}$ is a family of n convex sets in $\mathbb{R}^d$, with $n > d$. Then if every $d+1$ elements of $\mathcal{K}$ have a non-empty intersection,

$$\bigcap_{K_i\in\mathcal{K}} K_i \neq \emptyset$$

The next lemma is a useful result about convex sets:

Convex Intersections Lemma: For n convex sets to realize every possible intersection (i.e. $U_1 \cap ... \cap U_k \not\subset U_{k+1} \cup ... \cup U_n$, and similarly under any permutation of the variables) requires at least $n-1$ dimensions.

Proof: This can be realized by the symmetric placement of n $n-1$-spheres with their centers on the vertices of an $(n-1)$-simplex with equal radii that are large enough so that they all intersect, but smaller than the edge length of the simplex. (See figure 1 for an example with $n = 3$.) That it is impossible to realize this in fewer dimensions can be seen by induction on n, with base case $n = 2$, which is clear, since two sets need at least one dimension in order to have points outside their intersection, and the inductive step proceeds because each increase in n requires adding a set which has the same relationship to the other $n-1$ sets as they have to eachother, which requires an extra dimension.

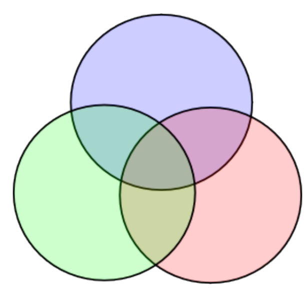

Figure 1: Intersection of 3 convex sets

The last result is due to Kalai ([7]):

At Most k Theorem: Suppose $\mathcal{K}$ is a family of convex sets in $\mathbb{R}^d$ for some $d \geq 1$, and $\mathcal{K}$ has an intersecting subfamily that misses at most $k \geq 0$ members of $\mathcal{K}$. Then the number of maximal intersecting subfamilies of $\mathcal{K}$ is at most

$$\binom{k+d}{d}$$

4 Embedding dimensions

Now that the principal definitions are behind us, we are interested in the following problem: Given $CF(J_C)$, (Or equivalently J_C or I_C or even $\Delta(C)$), and no knowledge about the stimulus space X, what can we ascertain about the minimal embedding dimension of the stimulus space, M? For example, it is proven in [1] by a straightforward application of Helly's theorem that:

Theorem 1: ([1]) If $\sigma \subset \{1, ..., n\}$ and the monomial $\prod_{i\in\sigma} x_i \in CF(J_C)$, then

$$M \geq |\sigma| - 1$$

Proof: For this monomial to appear in $CF(J_C)$, all intersections of $|\sigma| - 1$ sets which appear in it must be nonempty, but the intersections of all $|\sigma|$ must be empty. Helly's theorem forbids this in dimensions smaller than $|\sigma| - 1$.

However, it is in general quite difficult to infer the constraints on the dimension from the algebraic structure, as the pseudo-monomials interactions must generally be considered.

One difficulty is that adding pseudo-monomials to the canonical form does not always constrain the minimal embedding dimension more, in fact it may reduce it. For example, with $n = 4$, if $CF(J_C) = \emptyset$ the minimal embedding dimension is three (see later theorem.) However, if we add the relationships $x_1x_2, x_1x_3, ..., x_3x_4$, that is enforce that no sets intersect even pairwise, the minimal embedding dimension becomes 1.

For another example, suppose that $x_1x_2(1 - x_3)\cdots(1 - x_n) \in CF(J_C)$, and to make the problem even more general, assume that all the permutations of the variables in it also appear, e.g. $(1 - x_1)x_2(1 - x_3)x_4(1 - x_5)\cdots(1 - x_n) \in CF(J_C)$. This can still be embedded in $\mathbb{R}^2$, simply inscribe an n sided polygon in a circle, and omit the nth side piece from the nth set:

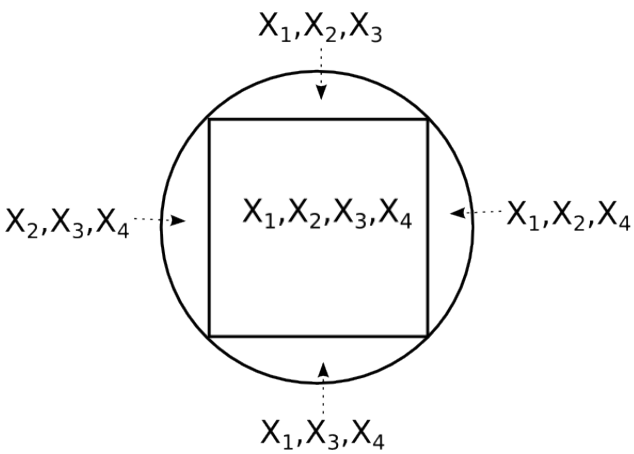

Figure 2: $x_1x_2(1 - x_3)(1 - x_4)$ and all permutations are in $CF(J_C)$

For yet another example, suppose that $x_1x_2\cdots x_{n-1}(1 - x_n) \in CF(J_C)$. This is realizable in 2 dimensions similarly to the previous problem, by inscribing an $n - 1$ sided polygon in a circle, having x_n be the polygon and each other x_i include the polygon and all but one of the outside chunks of the circle:

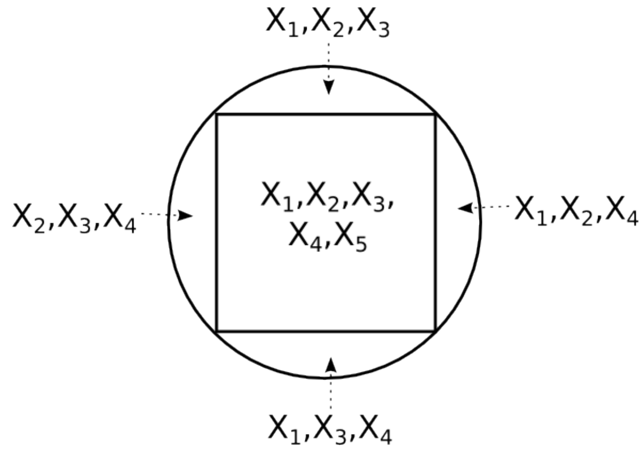

Figure 3: $x_1x_2\cdots x_{n-1}(1 - x_n) \in CF(J_C)$

In fact, for any pseudo-monomial which is not of the type I form, it is relatively easy to construct a collection of sets in $\mathbb{R}^2$ which will produce it alone. Thus again we must consider the interaction between the pseudomonomials, which is difficult. The next theorem does this by resorting to the simplicial complex structure, and using a result of Kalai.

Theorem 2: Let $\mathcal{P}_n$ be the set of all subsets of $\{1, ..., n\}$, and define

$$\begin{aligned} k &= \max\Big\{|Q| \ : \ Q \subset \mathcal{P}_n \text{ such that} \\ &\qquad \Big(\forall q_1 \in Q, \ \textstyle\prod_{i\in q_1} x_i \in CF(J_C)\Big) \text{ and} \\ &\qquad (\forall q_1, q_2 \in Q, q_1 \cap q_2 \neq \emptyset \Rightarrow q_1 = q_2)\Big\}. \end{aligned}$$

That is, k is the maximum number of empty intersections of receptive fields, each of which involves distinct receptive fields. Then we must have a minimal embedding dimension M such that:

$$\binom{k+M}{M} \geq i$$

where i is the number of intersections of order $n - k$,

$$i = |\{\sigma \in \Delta(C) : |\sigma| = n - k\}|.$$

Proof: If the conditions of the theorem hold, it is clear that the maximum intersecting families of $\mathcal{U}$ must miss at least k sets, since otherwise, by the pigeonhole principle our intersection would contain one of the k families with an empty intersection. Now, let $i =$

$|\{\sigma \in \Delta(C) : |\sigma| = n-k\}|$. Since the sets of cardinality $n-k$ are guaranteed to be maximal if k sets are missed in the maximal intersections by the At Most k Theorem, this provides the bound on M that

$$\binom{k+M}{M} \geq i$$

This theorem was chosen for its greater utility in the general situations where there will be many small pairwise intersections which are empty (for example, nerves on the skin which are mostly not adjacent), hence Theorem 1 will not be terribly useful as no high $|\sigma|$ terms will appear, since the lower intersections are empty. This theorem exploits these smaller empty intersections more efficiently, but it will only be useful if there are many maximal intersecting sets. For example, if $n = 4$, and x_1x_2 and x_3x_4 are the only Type I pseudo-monomials which appear in $CF(J_C)$, then this theorem implies that $M \geq 2$, whereas theorem 1 would only imply that $M \geq 1$.

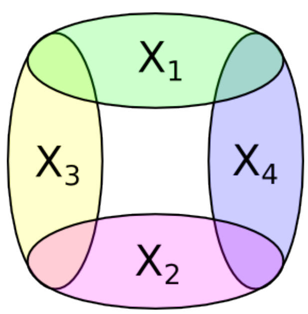

Figure 4: $CF(J_C) = \{x_1x_2, x_3x_4\}$

However, the theorem does provide a rather weak bound, due to the binomial coefficient on the right hand side, which grows quite rapidly. In addition, k can be at most $\lfloor n/2 \rfloor$, so this still requires some sufficiently large intersections when n is large, so it is not ideal.

The next theorem reverses the situation, it relies on the terms which do not appear in $CF(J_C)$, hence it may be more general in some cases than the previous ones, which relied on the appearance of specific terms in $CF(J_C)$.

Theorem 3: Suppose $\exists\gamma \subset \{1, ..., n\}$ such that $\forall \sigma, \tau \subset \gamma$ with $\sigma \cap \tau = \emptyset$,

$$\prod_{i \in \sigma} x_i \prod_{j \in \tau} (1 - x_j) \notin CF(J_C),$$

Then $M \geq |\gamma| - 1$.

Proof: The conditions on γ mean precisely that no relationship among the sets in γ appears. This means that every one of the $2^{|\gamma|}$ possible intersections of the $U_i, i \in \gamma$ must appear uniquely in the stimulus space, hence by the convex intersections lemma:

$$M \geq |\gamma| - 1$$

5 Conclusions & Looking ahead

In this paper we explained the general idea of the neural ring and ideal as developed in [1], and used provided two additional bounds on the minimal stimulus space dimension for a given neural code, which are useful in different cases than the theorem proven in [1]. Thus we are slowly chipping away at the problem of the embedding dimensions, but a general solution to this problem seems difficult and elusive. We provided some examples illustrating the difficulty therein, in particular we cannot even assume that adding further elements to $CF(J_C)$ will actually further restrict the dimension, which makes constructing general results difficult.

Further, some of the material cited in this paper, such as Eckhoff's conditions, we have not yet used to construct any bounds. We believe that these could be used, since in [6] Wegner proves that d-representible simplicial complexes are d-collapsible, hence Eckhoff's conditions indeed apply to our situation. The difficulty will lie in determining how the algebraic structures of the neural ideal constrain the simplicial complexes which Kalai's work uses. These results would be an interesting direction for further research.

Another possible direction is to exploit further geometric results. We have conjectured, but have not yet successfully proven the following:

Conjecture 1: Let $\sigma \subset \{1, ..., m\}$, and suppose that

$$\{(1 - x_i) \prod_{j \in \sigma, j \neq i} x_j \mid i \in \sigma\} \subset CF(J_C).$$

Then

$$M \geq |\sigma| - 4$$

We have sketched a proof for this based on the geometric problem of arranging the order $|\sigma| - 2$ intersections of the sets about the intersection of all of them, which exploits a version of Helly's theorem on the sphere. However, there is some difficulty with formalizing the geometric notions which we have been unable to overcome thus far. Regardless, this and other results of this type may be achievable by exploiting various notions from convex geometry.

Thus, although we have made some minor progress on the problem, there are many potential future areas for investigation on this subject. We suspect that further development will be quite rapid, due to the multifarious available approaches, especially if it is driven by the application of neural ring techniques to real neuroscience problems.

6 References

1. The neural ring: An algebraic tool for analyzing the intrinsic structure of neural codes, C. Curto, V. Itskov, A. Veliz-Cuba, & N. Youngs, 2013

2. Helly's theorem and its relatives, L. Danzer, B. Grünbaum, & V. Klee, 1963

3. Helly-Type Theorems and Geometric Transversals, R. Wenger, The Handbook of Discrete and Computational Geometry, 2004

4. Characterization of f-vectors of families of convex sets in $\mathbb{R}^d$. I: Necessity of Eckhoff's Conditions, G. Kalai, 1984

5. Characterization of f-vectors of families of convex sets in $\mathbb{R}^d$. II: Sufficiency of Eckhoff's Conditions, G. Kalai, 1986

6. d-Collapsing and nerves of families of convex sets, G. Wegner, 1975

7. Intersection Patterns of Convex Sets, G. Kalai, 1984