Population Stimulus-Response Functions

M. Sahani and J. F. Linden
Keck Center for Integrative Neuroscience
University of Califonia
513 Parnassus Ave
San Francsico, CA94143-0732

Abstract

Stimulus-response function models are typically estimated from sensory data one cell at a time. Population response properties can then be extracted from these cell-by-cell models, but doing so allows errors in the single cell characterizations to propagate. Moreover, these descriptions of population properties are often quite informal. Here, we attempt to obtain a systematic characterization of the population encoding properties directly from stimulus-response data collected for a group of similar cells from the same cortical area. We obtain (1) the stimulus subspace encoded by the ensemble of cells and (2) a sparse basis for this subspace, such that the linear receptive fields of the cells can each be described by a small number of basis vectors. These basis vectors thus provide a terse quantitative summary of the population encoding.

Extended Summary

A prerequisite for quantitative understanding of perceptual systems is the accurate modelling of sensory neuronal responses. Data for the construction of such models is often collected by exposing a subject to a long, continuously varying stimulus and measuring the activity of one or more sensory neurons. A number of techniques, "reverse correlation" being prominent among them, can then be used to estimate the sensory transformation that takes this stimulus to the measured neural response.

Typically, such response models are estimated from sensory data one cell at a time. However, particularly in mammalian cortex, sensory information appears to be conveyed by groups of neurons signalling in concert (in the form of "population codes"). In the conventional approach, the response properties of this population are extracted at a second stage, based on the cell-by-cell models. As a result, any errors in the characterization of the single cell models are allowed to propagate and may significantly affect the population description. Moreover, current descriptions of these properties are usually not systematic, being based on informal observations of receptive field shapes for individual cells, or on arbitrary classification of response features.

Here, we attempt to obtain a systematic characterization of the population encoding properties directly from the stimulus-response data for a group of similar cells from the same cortical area (not necessarily collected simultaneously). As in reverse-correlation studies, we assume that the cell responses are linear functions of some known aspect of the stimulus. We then characterize the population by describing (1) the stimulus subspace encoded by the ensemble of cells and (2) a sparse basis for this subspace, such that the linear receptive fields of each cell can be described by a small number of basis vectors. The net effect is similar in spirit to an under-complete independent components analysis (ICA) treatment of the linear "receptive fields" of the cells. The critical difference is with regard to the noise model implicit in such an analysis. By estimating the basis vectors with direct reference to the stimulus-response data rather than individually estimated models, we can shape the errors made in the individual cases to conform to match the discovered population description.

Our approach rests on the Bayesian evidence optimization framework for estimating single-cell response properties that we described briefly at the last CNS meeting, and in greater detail elsewhere.

In the most successful version of the this single-cell approach, the stimulus was first re-expressed in a new basis composed of smooth (Gaussian-shaped) basis vectors, the smoothness scales being chosen so as to optimize the Bayesian evidence (also called the marginal likelihood) of a hierarchical normal regression model. Once this optimal smoothed basis had been found, a second optimization of the evidence was used to find a sparse subset of the smooth basis vectors from which the linear response function could be constructed. This scheme was successful in the sense that it predicted the responses of the neuron to novel stimuli more accurately than competing approaches. However, both optimizations described were performed separately for each recording in the data base.

In our present work we have adapted this scheme so that the basis vectors chosen are shared between all recordings, although the linear response function that describes each recording is composed of a potentially different sparse subset of the basis vectors. Both the basis set and the individual sparse combinations are simultaneously optimized with respect to the *joint* evidence.

We have applied the techniques we develop to recordings from the thalamo-recipient layers of mouse cortical area A1, looking for response functions that are linear in the spectrogram of the sound stimulus. (In work presented at CNS last year, we showed that such models account for somewhere between 20 and 50% of the reproducible response power in rodent primary auditory cortex). The statements about predictive performance of the techniques made below pertain to this data set.

If the shared population basis vectors are collected into the columns of a matrix R, the evidence optimization proceeds by alternating updates to the sparse basis set for each recording (the form of which is identical to that in the previous work) and an update to the basis vectors, which is given by

$$R^{\text{new}} = \left(\sum_n \mathbf{w}_n^{\text{ML}} \langle \mathbf{w}_n^{\mathsf{T}} \rangle / \sigma_n^2 \right) \left(\sum_n \langle \mathbf{w}_n \mathbf{w}_n^{\mathsf{T}} \rangle / \sigma_n^2 \right)^{-1}$$

where the vector $\mathbf{w}_n$ indicates the current estimate of the linear response function for the nth recording, $\mathbf{w}_n^{\text{ML}}$ gives the corresponding maximum-likelihood (ML) estimate, σ_n^2 is the estimated noise level in the same recording and the expectations are taken over the current posterior distribution on $\mathbf{w}_n$. As can be seen, this is equivalent to a linear regression from the posterior weight distribution under the current probabilistic model to the ML weight vectors.

In the earlier work, the basis vectors were constrained to be Gaussian in shape, while here they are free to adopt any form. The resulting basis set is not necessarily orthogonal, but provides the preferred basis for sparsification of the linear response function weights. In practice, however, we find that for smaller groups of recordings this procedure is likely to over-fit and produce basis vectors that are tuned to the peculiarities (and noise) of the few recordings available. This is reflected in inferior predictions when compared to the best single-recording approach described above.

The scheme can be improved by adding a prior expectation that the basis vectors to be discovered will be smooth (although not constrained to Gaussian or any other form). If this prior expectation is expressed by a covariance matrix C, we obtain the following new update equation for R:

$$R \left(\sum_n \langle \mathbf{w}_n \mathbf{w}_n^{\mathsf{T}} \rangle / \sigma^2 \right) + (XX^{\mathsf{T}})^{-1} C^{-1} R = \left(\sum_n \mathbf{w}_n^{\text{ML}} \langle \mathbf{w}_n^{\mathsf{T}} \rangle / \sigma_n^2 \right)$$

where (XX^{T}) is the autocorrelation matrix of the stimulus. (This equation, which is a form of algebraic Riccati equation, is computationally impractical to solve algebraically, but can be solved efficiently by numerical techniques.) On our sample data set, predictions based on a smoothed basis obtained in this way are at least as accurate as can be obtained by best cell-by-cell model.