The Population Vector Could Implement Approximately Bayesian Inference

J. Orellana
R.E. Kass

Abstract

The brain uses populations of spiking neurons to encode, communicate, and combine sources of information, we are interested in how this process might be optimal as specified by Bayesian inference. Previous work from Kording and Wolpert [1] showed that performance of a sensorimotor task was consistent with optimal combination of sensory information and training knowledge; Ma et al. [2] also proposed a neural modeling framework according to which Bayesian inferences could be computed.These works focused on the form in which inputs were combined to produce the posterior mean and variance. We show that population vectors based on point process inputs combine evidence in a form that closely resembles Bayesian inference, with each input spike carrying information about the tuning of the input neuron. We investigated the performance of population vector-based inference with various tuning functions. We show that while its performance is exactly Bayesian for von Mises tuning functions, it remains approximately Bayesian for many other cases.

1 Introduction

There is considerable interest in understanding how the brain might use populations of spiking neurons to encode and communicate probability, and to combine sources of information in an optimal, or nearly optimal way, as specified by Bayesian inference. A useful review of the literature is provided by Beck et al. [3]. In neural encoding terms, a population represents information about a stimulus or behavioral feature using the simultaneous activity of a population of spiking neurons that are sensitive to that feature [4]. Far from being deterministic, the neural response for the same action or stimulus varies from trial to trial. This suggests that the brain might encode features as probability distributions (Eq. 3). For example, for a *center-out* reach action, the population code might represent a probability distribution with a central directional tendency μ_θ, and a measure of precision κ.

2 Population code

In this section we define a probabilistic model of population code (Eq. 3). First, let us suppose that spikes from each neuron i, within a population of N neurons follow independent point processes $r = \{r_i\}_{i=1,\dots,N}$ (Eq. 1, also see the inset panel of Figure 1). The mean response $f_i(\theta)$ depends on $\theta - \theta_{PDi}$, where θ is the intended direction of reach, and θ_{PDi} is the preferred direction for neuron i.

$$P(r_i|\theta) = \frac{exp\{-f_i(\theta)\}f_i(\theta)^{r_i}}{r_i!} \tag{1}$$

Next, we define f_i as a von Mises function (Eq. 2), which defines an exponential family on the unit circle, analogous to the normal distribution on the real line. Where A_i and B_i are constants

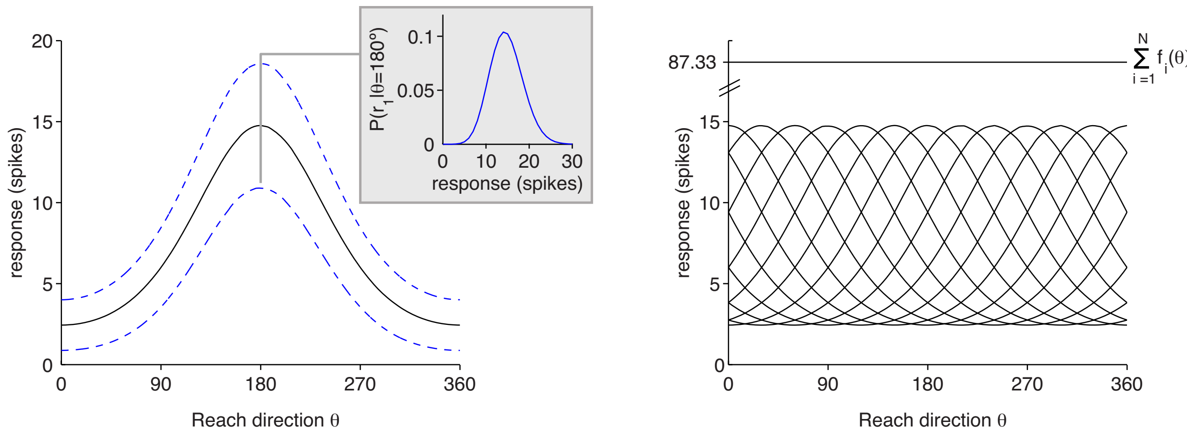

Figure 1: Encoding reach direction. The response distribution for one neuron with preferred direction of 180° is shown on the left panel (width at half amplitude= 133°). The black solid line indicates the mean, and the blue dashed lines are ± one standard deviation. The gray inset shows the poisson distribution for the neuron's response given the preferred direction stimulus. The right panel shows the mean response for a population of 12 neurons with equal precision and preferred directions spaced at 30°

representing the i_{th} neuron's amplitude and precision, respectively. High precision indicates narrow tuning for a particular preferred stimulus θ_{PDi}.

$$f_i(\theta) = A_i exp\{B_i \cos(\theta - \theta_{PDi})\} \tag{2}$$

The parameters θ, θ_{PDi}, and μ_θ are directional values; for a two-dimensional workspace, they can be conveniently expressed in circular angles $[0°, 360°]$ with $0°$ being equivalent to $360°$.
Figure 1 (left) shows the response vs. reach direction for a neuron with preferred direction of 180°. The noise observed in experimental recordings is typically approximated by a poisson distribution [2], such that the variance of the response is dependent on the direction stimuli with a variance to mean ratio of 1. Tuning curves for a population of N=12 neurons with equal precision ($B_i = B$) are shown on the right panel of Figure 1, the preferred directions are spaced by 30°. If we assume that every neuron responds independently, the population response distribution becomes the product of the individual neuron response distributions as shown in equation 3. We maintain the assumption of independence for mathematical simplicity, although experimental evidence shows that neural populations do exhibit correlations in firing rate.

$$P(r|\theta) = \prod_{i=1}^{N} P(r_i|\theta) \tag{3}$$

Now we discuss two ways of computing estimates of the intended direction stimulus from the population response: Bayesian inference and Population Vector.

3 Bayesian Inference

Bayesian decoders use Bayes' theorem to produce a posterior probability of the intended direction stimulus given the response:

$$P(\theta|r) = \frac{P(r|\theta)P(\theta)}{P(r)} \tag{4}$$

Where $P(r|\theta)$ and $P(\theta)$ are the likelihood function and prior distribution of the stimulus respectively and $P(r)$ is a normalizing constant. We obtained an expression for the posterior distribution in cartesian coordinates by assuming a uniform prior, which we will revisit in a later section, and combining equations 1-3.

$$P(\theta|r) \propto \prod_{i=1}^{N} P(r_i|\theta) \tag{5}$$

$$P(\theta|r) \propto \left(\prod_{i=1}^{N} \frac{1}{r_i!}\right)\left(exp\{-\sum_{i=1}^{N} f_i(\theta)\}\right)\left(exp\{\sum_{i=1}^{N} r_i \log(A_i)\}\right)\left(exp\{\sum_{i=1}^{N} r_i B_i \cos(\theta-\theta_{PDi})\}\right) \tag{6}$$

$\sum_{i=1}^{N} f_i(\theta)$ is constant over θ when the population has a uniformly dense distribution of preferred directions as shown on Figure 1B. Hence equation 6 is simply an non-normalized von Mises distribution governed by the last term on the right, $exp\{\sum_{i=1}^{N} r_i B_i \cos(\theta - \theta_{PDi})$. Let $\bar{S}_b = N^{-1}\sum_{i=1}^{N} r_i B_i \sin(\theta_{PDi})$, and $\bar{C}_b = N^{-1}\sum_{i=1}^{N} r_i B_i \cos(\theta_{PDi})$. We can define the concentration parameter as $\kappa^2 = \bar{S}_b^2 + \bar{C}_b^2$, the central tendency as $\hat{\mu}_b = \arctan(\bar{S}_b/\bar{C}_b)$, and the normalizing constant as $\tilde{A} = [2\pi I_0(\kappa)]^{-1}$ with I_0 being the modified Bessel function of order zero. Thus the posterior expression becomes:

$$P(\theta|r) = \tilde{A} exp\{\kappa \cos(\theta - \hat{\mu}_b)\}\} \tag{7}$$

A 95% credible interval for the central tendency can be calculated directly from the von Mises probability distribution in equation 7 such that:

$$P(\hat{\mu}_b - \theta_b^* \leq \mu_b \leq \hat{\mu}_b + \theta_b^*) = 0.95 \tag{8}$$

Thus the angular size of the credible interval for the decoded stimulus is given as $L_b = 2\theta_b^*$

4 Population Vector

An alternative to Bayesian inference is a population vector estimate, which is a simple way to compute an estimate of the stimulus from the population response [5]. The direction stimulus estimate is an average of preferred directions weighted only by the activity of each corresponding neuron. Let $\bar{S}_{pv} = N^{-1}\sum_{i=1}^{N} r_i \sin(\theta_{PDi})$, and $\bar{C}_{pv} = N^{-1}\sum_{i=1}^{N} r_i \cos(\theta_{PDi})$. Note the absence of B_i compared to Bayesian Inference. The resultant magnitude and direction are given by $\bar{R}^2 = \bar{S}_{pv}^2 + \bar{C}_{pv}^2$, and $\hat{\mu}_{pv} = \arctan(\bar{S}_{pv}/\bar{C}_{pv})$ respectively.

We can also think of the population vector as an estimate resulting from every spike r_i carrying directional information from its emitting neuron's preferred direction θ_{PDi}. With this in mind, we consider spikes emitted by the population as samples from a circular random variable with a well defined mean direction μ_{pv}. Let $\alpha_2 = N^{-1}\sum_{j=1}^{N} r_i \cos 2(\theta_{PDi} - \hat{\mu}_{pv})$. We use the Circular Central Limit Theorem [6] to obtain an approximate 95% confidence interval for μ_{pv} as $\hat{\mu}_{pv} \pm \sin^{-1}(1.96\hat{\sigma}_{pv})$ with $\hat{\sigma}_{pv} = \{(1-\alpha_2)/(2M\bar{R}^2)\}^{1/2}$ as the circular standard error. The angular size of the confidence interval for the decoded stimulus is given as $L_{pv} = 2\sin^{-1}(1.96\hat{\sigma}_{pv})$.

5 Comparing population estimates

In this section we compare the uncertainty associated with estimating the direction of reach using both Bayesian Inference and Population Vector. Consider the population response to one instance in which the intended reach direction is 180°. Figure 2 (left) shows the population response plotted against the preferred direction of each neuron. The estimates of the direction of reach are the maximum likelihood $\hat{\mu}_b$ (posterior mean) for Bayesian Inference, and the activity-weighted average direction $\hat{\mu}_{pv}$ for population vector. Note that these two estimates are equal for the special case of uniform precision encoding in the population. That is if $B_i = B$ (see Figure 1) then it follows that:

$$\begin{aligned}
\hat{\mu}_b &= \arctan(\bar{S}_b/\bar{C}_b) \\
&= \arctan(\frac{N^{-1}B\sum_{i=1}^{N} r_i \sin(\theta_{PDi})}{N^{-1}B\sum_{i=1}^{N} r_i \cos(\theta_{PDi})}) \\
&= \arctan(\bar{S}_{pv}/\bar{C}_{pv}) \\
\hat{\mu}_b &= \hat{\mu}_{pv}
\end{aligned} \tag{9}$$

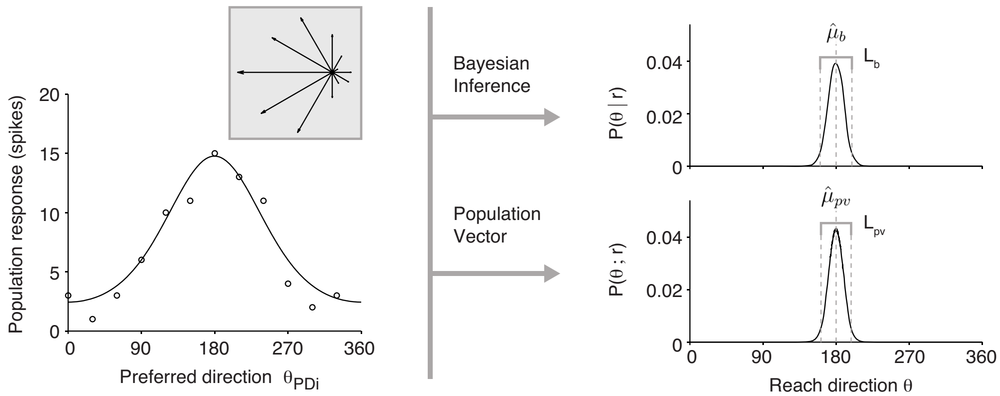

Figure 2: Computing estimates of the stimulus from the population response to an intended reach direction of 180° using Bayesian inference and Population Vector. Left: Population response plotted against the preferred direction of each neuron and shown in cartesian coordinates (inset). Right: Posterior probability distribution of the stimulus given the response using Bayesian inference (top), and probability distribution of the stimulus using Population Vector and circular Central Limit Theorem (bottom).

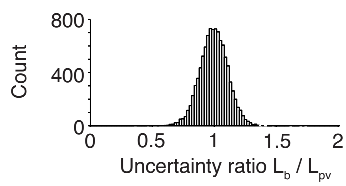

Figure 3: Distribution of the uncertainty ratio of credible to confidence interval for 10,000 repetitions for the population of neurons shown in Figure 1 and a stimulus of 180°

Yet the uncertainty associated with each estimate is similar but not necessarily equal. Figure 2 (Right) shows the respective probability densities associated with each estimate. The credible interval of size L_b was obtained by applying Bayesian Inference under the assumption of a uniform prior distribution. On the contrary, the confidence interval of size L_{pv} was obtained using the Central Limit Theorem. When repeatedly computing the ratio of credible interval size to confidence interval size we observe that the distribution is centered at 1 with a standard deviation of 0.112 (Figure 3). This suggests that although not exactly equal, the uncertainty of L_{pv} tends to be approximately equal to that of L_b.

References

[1] Krding, Konrad P., and Daniel M. Wolpert. ”Bayesian integration in sensorimotor learning.” Nature 427.6971 (2004): 244-247.

[2] Ma, Wei Ji, et al. ”Bayesian inference with probabilistic population codes.” Nature neuroscience 9.11 (2006): 1432-1438.

[3] Beck, Jeffrey M., et al. ”Probabilistic population codes for Bayesian decision making.” Neuron 60.6 (2008): 1142-1152.

[4] Ma, W. J., and A. Pouget. ”Population Codes: theoretic aspects.” Encyclopedia of neuroscience 7 (2009): 749-755.

[5] Georgopoulos, Apostolos P., Ronald E. Kettner, and Andrew B. Schwartz. ”Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. Coding of the direction of movement by a neuronal population.” The Journal of Neuroscience 8.8 (1988): 2928-2937.

[6] Fisher, Nicholas I., and Toby Lewis. ”Estimating the common mean direction of several circular or spherical distributions with differing dispersions.” Biometrika 70.2 (1983): 333-341.