Probabilistic Theories of the Visual Cortex

A.L. Yuille and D. Kersten

Abstract

PROBABILITY MODELS OF GROUPS OF NEURONS

PURPOSE: TO DESCRIBE MODELS FOR INTERACTIONS NEURONS. TO DISCUSS DYNAMICS – GIBBS SAMPLING AND MFT (MENTION BP ALSO). TO SHOW THAT DYNAMIC "NEURAL MODELS" CAN BE DERIVED FROM MRF'S BY MFT. MAKES FUNDAMENTAL LINK BETWEEN COMPUTATIONAL MODELS – MRF'S - AND NEURAL MODELS.

This section describes models of groups of neurons and neural algorithms. We relate these neural algorithms to inference algorithms for computational models (e.g., to perform inference for visual tasks such as segmenting image and estimating depth – discussed in later lectures). These models can also be used for estimating the activity of populations of neurons. Using multi-electrode recording to measure the activity of neurons in closely similar positions. These experiments – e.g, on interpreting the activity in the monkey motor planning cortex enabling the monkey to control a cursor – suggest that is possible to partially decode the activity of groups of neurons. This requires fitting parametric models to the data (see next lecture).

A. Boltxmann Machines, Markov Random Field Models, and Gibbs Sampling

How to extend the simplified neural model, see previous lecture, to a population of neurons? The simplest way is by defining a Markov Random Field (MRF). This represents the activity $\vec{o} = (o_1, ..., o_n)$ of a group of n neurons by a probability distribution:

$$P(\vec{x}|\vec{I}) = \frac{1}{Z} \exp\{\sum_{ij} T_{ij} x_i I_j + (1/2) \sum_{ij} \theta_{ij} x_i x_j\}. \tag{1}$$

This model includes a term $\sum_{ij} T_{ij} x_i x_j$ which specifies interactions between the neurons. If this term was not present, then this distribution would reduce to independent copies of our simple model:

$$P(\vec{x}|\vec{I}) = \frac{1}{Z_T} \exp\{\sum_{ij} T_{ij} x_i I_j\} = \prod_{i=1}^{n} P(x_i|\vec{I}). \tag{2}$$

Alternatively, if we consider a single neuron i and fix the activity of the other neurons, then the conditional probability of its response is given by:

$$P(x_i|\vec{I}, \vec{x}_{/i}) = \frac{1}{Z_i} \exp\{x_i(\sum_{j} T_{ij} I_j + \sum_{j \neq i} \theta_{ij} x_j)\}. \tag{3}$$

Hence we can think of other neurons as giving additional inputs (plus the inputs I_i).

The probability distribution in equation (1) specifies a Boltzmann Machine (though this requires the learning algorithm – see next lecture). It is also a Markov Random Field (MRF). It enables interactions between neurons and will be used in later lectures to model non-classical receptive fields and a range of visual perceptual phenomena (including segmentation and binocular stereo).

Now consider the stochastic neural model (previous lecture) and update a single neuron conditioned on all the others. This reduces to sampling from the conditional distribution in equation (**??**).

This has an interpretation in terms of MRFs. If we select each neuron at random and update it by sampling from equation (**??**), then this is the Gibbs sampling algorithm. On a serial computer, *Gibbs sampling* proceeds by sampling the state of each neuron x_i conditioned on the states $\vec{x}_{/i}$ of the other neurons. This sampling is done from the conditional distribution $P(x_i|\vec{I}, \vec{x}_{/i})$ specified by equation (3). By the theory of MCMC, repeated sampling for different neurons will converge to samples from the full distribution $P(\vec{x}|\vec{I})$ which takes into account the interactions between the neurons. On a parallel computer, or to specify the dynamics of a parallel neural network model, the sampling is done simultaneously (or with a small time delay, as described below).

This enables us to do inference on MRFs (e.g., segmenting images, estimating depth). Hence these stochastic neural algorithms can be used as ways to implement computational model by neurons.

B. Dynamic Neural Models and mean Field Theory

A second class of neural model are described by differential equations (Hopfield, Grossberg, Cowan) [7]) These correspond to *mean field theory dynamics* on MRFs. The mean field approach is an approximation developed in statistics physics (refs!!) whose aim is to specify dynamics which converges to an approximation to the expected states $\sum_{\vec{x}} \vec{x} P(\vec{x}|\vec{I})$. It has been popularized as variational methods (refs!!) where it is formulated to find the factorized distribution $Q(\vec{x}) = \prod_{i=1}^{n} q_i(x_i)$ which is most similar to $P(\vec{x}|\vec{I})$. This is formulated to minimize the Kullback-Leibler divergence:

$$F(Q) = \sum_{\vec{x}} Q(\vec{x}) \log \frac{Q(\vec{x})}{P(\vec{x}|\vec{I})} = -\sum_{ij} T_{ij} I_j q_i - (1/2) \sum_{ij} \theta_{ij} q_i q_j + \sum_i \{q_i \log q_i + (1-q_i) \log(1-q_i)\} + \log Z, \tag{4}$$

where $q_i = \sum_{x_i} x_i q_i(x_i) = q_i(x_i = 1)$.

The idea is that we can compute/infer any property we want from the best factorized distribution $Q^* = \arg\min F(Q)$. This reduces inference to finding Q^*. But usually $F(Q)$ is non-convex and so no algorithms are guaranteed to find the global minimum. But instead we can specify algorithms that can converge to local minima of $F(Q)$.

Dynamic steepest descent (modified) is an algorithm that will converge to a local minimum of $F(Q)$:

$$\frac{dq_i}{dt} = -q_i(1 - q_i)\frac{\partial F(Q)}{\partial q_i}. \tag{5}$$

These dynamics are guaranteed to decrease the quantity $F(Q)$ monotonically with time t (because $dF/dt = -\sum_i (\partial F/\partial q_i)(\partial q_i/\partial t) = -\sum_i q_i(1 - q_i)(\partial F/\partial q_i)^2$). The term $q_i(1 - q_i)$ modifies steepest descent but does not affect convergence.

To relate this mean field dynamics directly to dynamical system models based of neurons. We do this by introducing a new variable $u_i = \log q_i/(1 - q_i)$, which implies that $q_i = \sigma(u_i)$. Noting that $\partial F/\partial q_i = -\sum_j T_{ij} I_j - \sum_j \theta_{ij} q_j + \log q_i/(1 - q_i)$, $u_i = \log q_i/(1 - q_i)$, and $dq_i/q_i(1 - q_i) = du_i$ we obtain:

$$\frac{du_i}{dt} = -u_i + \sum_j T_{ij} I_j + \sum_j \theta_{ij} q_j. \tag{6}$$

These "neural dynamics" sum the inputs on the dendridic tree to estimate activity u_i at the soma, which determines the output $q_i = \sigma(u_i)$ (roughly the probability that the cell will fire). Observe that this is similar to taking the expectation of the Gibbs sampling rule. They relate to a class of neural network models – Cowan, Grossberg, [1].

C. Discrete Update MFT

An alternative MFT algorithm shows a relationship between MFT and Gibbs samling (i.e. stochastic models of neurons). This is used to justify the use of MRF to approximate the activity of populations of neurons.

An iterative steepest descent algorithm is guaranteed to monotonically decrease $F(Q)$ (subject to some conditions of θ_{ij}):

$$q_i^{t+1} = \frac{\exp\{\{\sum_j T_{ij}I_j + \sum_j \theta_{ij}Q_j\}\}}{1 + \exp\{\{\sum_j T_{ij}I_j + \sum_j \theta_{ij}q_j\}\}}. \tag{7}$$

This can be obtained as an approximation to the Gibbs sampler. To do this, set $u_i = \sum_{x_i=0}^{1} x_i P(\vec{x})$. Then taking expectations of the Gibbs sampling gives:

$$\frac{du_i}{dt} = \sigma(\sum_j T_{ij}I_j + \sum_j \theta_{ij}u_j) \tag{8}$$

This will converge to a local minimum of the free energy:

$$F = \sum_{ij} T_{ij}u_iI_j + \sum_{ij} \theta_{ij}u_iu_j + \sum_i \{u_i \log u_i + (1 - u_i)\log(1 - u_i)\} \tag{9}$$

This relation between MFT and Gibbs sampling was established in Statistical Physics. Note that MFT relates to belief propagation and the Bethe Free Energy ($F(Q)$ is a simply free energy). Belief propagation can also be related to a form of Gibbs sampling (Rosen-Zvi, Jordan, Yuille).

D. Potts Models

In this section we have described MFR model where the state variables x_i take values $0, 1$. MRF's can be extended to distributions where the x_i take a larger set of values. These are called Potts models. They are used for visual computational tasks (e.g., binocular stereo). They can be implemented by groups of neurons, where different neurons correspond to different states of x_i. This will be discussed in later sections.

E. Generalized Linear Models (GLM)

PURPOSE: TO DISCUSS HOW MODIFICATIONS OF THE MRF'S CAN BE USED TO DESCRIBE GROUPS OF NEURONS AND TO ANALYSE THE DATA FROM MULTI-ELECTRODE RECORDINGS.

These neural are over-simplified in many respects. This section discusses more realistic models.

In the last ten years there have been interesting attempts to model the properties of groups of neurons and relate them to the multi-electrode recordings. This takes models which are more complex that those described above (in terms of dynamics) and estimates their parameters to best fit the neural activity. These class of models include Poisson models for neurons and extensions such as Generalized Linear Models (GLM), see Brown *et al.* [2],[3], Pillow *et al.* [5], [6]. Background references include Dayan and Abbott [4].

$$P(n|\lambda) = \frac{(\lambda\Delta)^n}{n} \exp\{-\lambda\Delta\}. \tag{10}$$

For a spike train $N = \{n_t\}$ the firing rate depends on parameters $\theta = \{\lambda_t\}$ which can changes over time (e.g., as the input to the neuron changes). This gives a probability model:

$$P(N|\theta) = \prod_t \frac{(\lambda_t\Delta)^{n_t}}{n_t} \exp\{-\lambda_t\Delta\}. \tag{11}$$

(Note: add references to Pillow et al, T-S Lee at al.).

A simple task is to estimate the parameters θ of the model from a set of observed spikes N. This can be expressed as maximum likelihood estimation: $\theta^* = \arg\max P(N|\theta)$. For this simple model, we calculate $\log P(N|\theta) = \sum_t n_t \log \lambda_t + \sum_t n_t \log \Delta - \sum_t \log n_t - \Delta \sum_t \lambda_t$, differentiate with respect to λ_t, and obtain the solution $\hat{\lambda}_t = y_t/\Delta$.

Next we consider a more realistic case where we seek to estimate receptive field properties of neurons and their refractoriness. The parameters of the model are $\theta = (\vec{k}, \vec{h}, \mu)$, where $\vec{k}$ denotes the receptive field of the neuron, $\vec{h}$ models spike history dynamics (refractoriness, bursting), and μ is a constant bias. We predict the response of the neuron at time period t dependent on its inputs $\vec{x}_t = (x_{t-1}, x_{t-2}, ...)$ and previous spike counts $\vec{n}_t = (n_{t-1}, n_{t-2}, ...)$. We model the firing rate λ_t to depend on the input and previous time history by:

$$\lambda_t = \exp\{\vec{k} \cdot \vec{x}_t + \vec{h} \cdot \vec{n}_t + \mu\}. \tag{12}$$

Hence the firing rate of the neuron will be influenced by its inputs and its previous firing by amounts which are quantified by $\vec{k}$ and $\vec{h}$. The log-likelihood becomes:

$$\log P(N|\theta) = \sum_t n_t\{\vec{k} \cdot \vec{x}_t + \vec{h} \cdot \vec{n}_t + \mu\} - \Delta \sum_t \exp\{\vec{k} \cdot \vec{x}_t + \vec{h} \cdot \vec{n}_t + \mu\} + \text{Constant}. \tag{13}$$

This log-likelihood is a convex function of the parameters $\theta = (\vec{k}, \vec{h}, \mu)$ and can be solved by gradient descent methods.

We now consider networks of neurons. To enable the model to capture correlated activity, like synchrony, we allow spikes from one neuron to influence the firing rate of other neurons. This gives the firing rate of neuron i to be:

$$\lambda_i = \exp\{\vec{k}_i \cdot x + \sum_j \vec{h}_{ij} \cdot \vec{y}_j + \mu_i\}. \tag{14}$$

The term $\sum_j \vec{h}_{ij} \cdot \vec{y}_j$ includes input from other neurons ($j \neq i$), or self-interaction (i.e. refraction).

This model specifies the response of neurons as before $P(n^i|\lambda_i)$ is given by equation (10) for the ith neuron, but where the firing rate λ_i are specified as above.

References

[1] L. Abbott and T. Kepler. Model neurons: From hodgkin-huxley to hopfield. *Statistical mechanics of neural networks*, pages 5–18, 1990.

[2] E. Brown, D. Nguyen, L. Frank, and S. V. Wilson, M. *Proceedings of the National Academy of Sciences*, 98(21):12261, 2001.

[3] E. N. Brown, D. P. Nguyen, L. M. Frank, M. A. Wilson, and V. Solo. An analysis of neural receptive field plasticity by point process adaptive filtering. *Proceedings of the National Academy of Sciences of the United States of America*, 98(21):12261–12266, Oct. 2001.

[4] P. Dayan and L. Abbott. The MIT Press, Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems, 2001.

[5] J. Pillow, J. Shlens, L. Paninski, A. Sher, A. Litke, E. Chichilnisky, and E. Simoncelli. Spatiotemporal correlations and visual signaling in a complete neuronal population. *Nature*, 454:995–999, 2008.

[6] J. Shlens. *Notes on Generalized Linear Models of Neurons*. Systems Neurobiology Laboratory, Salk Insitute for Biological Studies, La Jolla, CA 92037, 2008.

[7] A. Yuille. Energy functions for early vision and analog networks. *Biological Cybernetics*, 61:115–123, 1987.