

Gatsby Computational Neuroscience Unit Combined Qualifying Examination

21 Jan 2016

There are 10 questions. Answer all of them, to the best of your ability. Each is worth 10 marks. No reference materials are allowed.

It should take no more than 6 hours

Good luck!

1. A neurophysiologist collects the activity of a large number of LIP neurons simultaneously on a very large number of trials in the classic fixed-duration “dots” task: i.e., a monkey is shown a field of noisy moving dots for a duration determined by the experimenter, following which he must indicate the predominant direction of motion to obtain some water.

Having collected her data, the physiologist constructs a linear choice “decoder”, using the population neural activity $\mathbf{n}(t)$ in each time bin t during dot presentation (each element of the vector $\mathbf{n}(t)$ is the spike count of one neuron in bin t) to obtain a time-varying “decision variable” $d(t)$ that gives the log-odds of the monkey choosing rightward motion at the end of the trial:

$$d(t) = \boldsymbol{\beta}^\top \mathbf{n}(t)$$

$$P(\text{choice} = \text{right}) = \frac{1}{1 + e^{-d(t)}}.$$

She estimates the linear parameters $\boldsymbol{\beta}$ by maximum-likelihood, treating each population vector $\mathbf{n}(t)$ as a separate input associated with observed choice (i.e., no time series model).

She then performs a new experiment in which she monitors $d(t)$ in real time (using the same neurons and same decoder weights) while the dots are being presented, and terminates the presentation of the dots early whenever $d(t)$ crosses a threshold value D . The monkey indicates his decision after the dots turn off, and the experimenter compares the fraction of rightward decisions on the early-termination trials to the value $\frac{1}{1+e^{-D}}$. She finds that her logistic “prediction” fits the monkey’s behaviour perfectly.

Suppose first that the monkey’s behaviour is optimal, and that the estimated $d(t)$ does in fact reflect a noiseless reconstruction of his internal belief (measured in log-odds) about the direction of the dots.

- (a) By considering the remaining sources of variability in the two experiments, explain why this second result is *inconsistent* with these assumptions of optimality.

Speculate on a possible form of sub-optimality in each of the:

- (b) animal’s behaviour; or
- (c) decoding process

that could lead to the result seen.

2. A population of simple cells with identical spatial receptive field centres and spatial frequency tuning have orientation tuning curves given by $f_i(\theta) = f(\theta - \theta_i)$ where θ_i is the preferred orientation of the i th neuron and $f()$ is a symmetric bump-shaped function.

Suppose that, when stimulated by a drifting grating with orientation θ , contrast a and spatial frequency ω the cells fire at mean rates given by $\mu_i = ac(\omega)f(\theta - \theta_i)$ (integrating over the grating phase), with correlated variability so that the covariance between the rates of neurons i and j is of the form: $\Sigma_{ij} = \sigma^2(a, \omega)\rho(\theta_i - \theta, \theta_j - \theta) + \alpha^2\delta_{ij}$ where δ_{ij} is the Kronecker delta.

Treating the population variability as Gaussian, show how the (Fisher) information about a , ω and θ depends on the correlation structure of noise in the observed rates. In particular, consider cases where:

- (a) $\rho() = \text{constant}$,
- (b) $\rho() \propto f_i(\theta)f_j(\theta)$, and
- (c) $\rho() \propto f'_i(\theta)f'_j(\theta)$.

Provide both formal and intuitive explanations for your claims: in particular, what aspects of the population response carry information about each stimulus parameter? Remember to consider both the marginal and joint information about the 3 parameters.

3. Spike trains $\{s_i(t)\}$ are recorded simultaneously from a group of sensory cortical neurons (indexed by i) responding to the same time varying stimulus $x(t)$. Using the sum-of-deltas representation of $\{s_i(t)\}$,

- (a) write an expression for the conditional intensity of the i th neuron in a conventional exponential-link “coupled-GLM” model (e.g. of the type studied by Pillow et al (2008) for the retina).

One alternative is to use a multivariate Hawkes process instead, with

$$\lambda_i(t) = \mu_i[x[0, t]](t) + \sum_j \int_0^t d\tau k_{ji}(t - \tau)s_j(\tau)$$

where $\mu_i[\cdot]$ is a functional transforming the stimulus history in the interval $[0, t)$ to a base rate for neuron i , and the terms $k_{ji}()$ reflect the spike-dependent self- and cross-coupling.

- (b) Compare the ability of the GLM and Hawkes approaches to model the phenomena we expect in neural data. Consider known biophysical properties of neurons, patterns of connectivity and stability.

Either class of model may be augmented by introducing an unobserved time-series $z(t)$ (possibly itself multivariate) which provides an additional “input” to the intensity function of each observed neuron. Suppose that $z(t)$ is Gaussian with Markov dependence (or, in continuous time, is the solution to a first-order SDE with Wiener-process fluctuations).

- (c) What form of covariance will this create in the outputs of each type of process considered above?
 (d) Again, discuss which might be a more appropriate model for a neural population.

4. consider a feedforward linear mapping from n -d input \mathbf{x} to m -d output \mathbf{y} via $m \times n$ weight matrix W , so $\mathbf{y} = W\mathbf{x}$. Williams’ symmetric error correcting learning rule is a form of predictive coding-based learning, suggesting that

$$\Delta W_{ij} \propto y_i(x_j - [U\mathbf{y}]_j)$$

where U is a $n \times m$ top-down matrix.

- write down the equation for the evolution of the weights, averaging over the statistics of the input patterns $\{\mathbf{x}\}$.
 - in the symmetrized version of the rule $U = W^T$. In this case, write down the condition that must be satisfied for W to stop changing.
 - writing $P = W^T W$, show that these equilibrium points satisfy $PR = RP = PRP$ for a symmetric matrix R (which you should define)
 - if the (orthonormal) eigenvectors $\{\mathbf{e}_\alpha\}$ and eigenvalues $\{\lambda_\alpha\}$ of R are such that $\{\lambda_\alpha\}$ are all different, show that P has the property that $P\mathbf{e}_\alpha = \mathbf{e}_\alpha$ or $P\mathbf{e}_\alpha = 0$.
 - interpret what this implies as an end-point of learning for the rule.
5. consider a reward machine that has two states: empty (0) and full (1), with a probability of transitioning from $0 \rightarrow 1$ of $p = 0.2$ per step. An agent can inspect the machine (at cost a); if it finds it to be full, then it gets a reward of $r = 1$ (and the machine goes back to state 0). Otherwise, the machine remains empty. If it is full and isn’t inspected, then it remains full.
- if the agent inspects with probability π , what is the long run probability that the machine is in state 0 or 1?
 - derive an expression for the inspection frequency π^* that optimizes the long run average net reward per step as a function of a . Sketch how this long run average reward varies with π for indicative values of a , both smaller and greater than 1.
 - if the agent instead inspects every n steps on a deterministic schedule, what is the long-run probability that it will find the machine full when it inspects?
 - will optimal deterministic inspection lead to greater, equal or lesser long run reward per step than optimal stochastic inspection. Justify your answer.
6. consider an axon growth cone growing in an environment in which the gradient of a guidance molecule points at angle ψ relative to the origin. The cone takes discrete steps of length 1 unit, starting from an anchoring point at the origin, and pointing at an angle $\theta(t)$. Axons like to grow straight, and so, ignoring noise, a reasonable model of the change in $\theta(t)$ over time is:

$$\theta(t+1) = \theta(t) + \alpha \angle(\phi(t), \theta(t)) + \beta \angle(\psi, \theta(t))$$

where $\angle(\alpha, \beta)$ denotes the signed angle between the unit vectors with angles α, β , and constrains the resultant angle to be between $-\pi$ and π , $\phi(t)$ is the angle joining the tip of the axon to the origin, $\alpha > 0$ quantifies the force of straightness (making θ align with ϕ), and $\beta > 0$ quantifies the force of the gradient (making θ align with ψ). Let $\psi = 0$.

- (a) if the axon starts off growing at $\theta(0) = \psi = 0$, write down a formula for $\theta(t)$ as a function of t .
 (b) take the case that the axon starts with $\theta(0) = \pi/2$, and in which the axon is quite long, so that $\psi(t)$ changes much more slowly than $\theta(t)$.
- at what value does $\theta(t)$ equilibrate (as a function of $\psi(t)$)?
 - assuming that equilibrium is maintained as $\phi(t)$ changes, sketch the form of the evolution of the axon

- and by writing down an equation for the evolution of $\phi(t)$ and doing a first order approximation, write down a closed-form expression for its long-run value. “long-run value” means “ $t \rightarrow \infty$ ”. at finite times you will get an equation you can’t solve.
 - if the axon can put down more anchor points as it grows, changing the origin to which $\phi(t)$ refers, need this form remain true? why?
7. Describe the sequence of events that lead to NMDA-mediated long-term potentiation. Are these events consistent with STDP (spike-time dependent plasticity)? Explain your reasoning.
8. There are now several neurons in the “IF” class: LIF, QIF and EIF. Write down the membrane potential equations for all three. Explain why they all describe type I neurons. For any one of the IF models (you choose), modify the dynamics so that the equations admit a limit cycle (without the membrane potential crossing threshold), potentially turning them into type II neurons. More specifically, consider new equations,

$$\tau_V \frac{dV}{dt} = f(V) + X \quad (1a)$$

$$\tau_X \frac{dX}{dt} = g(V, X) \quad (1b)$$

where $f(V)$ correspond to the IF model of your choice. Choose $g(V, X)$ so that the equation admits a limit cycle. You can demonstrate the existence of the limit cycle using nullclines and sketching trajectories; you don’t have to find the eigenvalues explicitly (although you can if you want to).

This was a badly worded question. Normally to show that a neuron is type II you need to show that it exhibits a subcritical bifurcation. Here you just need to show that a limit cycle exists. Importantly, the limit cycle should exist without the membrane going to infinity (QIF or EIF) or hitting threshold (LIF).

9. Consider binary synapses – synapses with two discrete states. For convenience we’ll let the two states be -1 and $+1$ (corresponding to depressed and potentiated, respectively). We’ll discretize time, and assume that on any time step the synapses change probabilistically according to

$$\begin{aligned} -1 &\rightarrow +1 && \text{probability } 1 - p \\ +1 &\rightarrow -1 && \text{probability } 1 - p. \end{aligned} \quad (2)$$

At $t = 0$, half the synapses are set to 1, and the other half are set to -1 ; after that, the synapses continue to evolve according to Eq. (2). The “signal,” $S(t)$, is

$$S(t) = \frac{1}{n} \sum_i w_i s_i(t) \quad (3)$$

where $s_i(t)$ is the state of synapse i (-1 or $+1$) at time t and $w_i = s_i(0)$. At $t = 0$, $S(t) = 1$, but it decays with time, eventually approaching 0 on average. Show that the mean and variance of $S(t)$ are given by

$$\langle S(t) \rangle = \frac{1}{n} \sum_i w_i \langle s_i(t) \rangle \quad (4)$$

$$\text{Var}[S(t)] = \frac{1}{n^2} \sum_i (1 - \langle s_i(t) \rangle^2). \quad (5)$$

Thus, if you know how $s_i(t)$ behaves on average, you know everything. Write down an explicit expression for $\langle s_i(t) \rangle$ versus t . Based on that expression, and the expressions for the mean and variance, how long can a “memory” of this type last?

10. Consider a very approximate model of a network of n of excitatory linear integrate and fire neurons,

$$\tau \frac{dV_i}{dt} = -V_i + w_i \bar{\nu} \quad (6)$$

where ν_i is the firing rate of neuron i , and $\bar{\nu}$ is the population averaged firing rate

$$\bar{\nu} \equiv \frac{1}{n} \sum_{i=1}^n \nu_i. \quad (7)$$

A spike is emitted when $V_i = V_{th}$, at which point the membrane potential is set to 0.

Derive an equation for the population averaged firing rate. Show that if all the weights are the same, $w_i = w$, then there is a solution if and only if $w > V_{th}\tau$. Is the solution stable or unstable?