Wednesday, February 20, 2019 2:00 PM

?? = unsure about what professor said
? = actual question/ point of interest
projecting a stimulus onto the axis, would tell you something abour r, but it would tell you less about the stimulus in the unscaled density curve.

- Tatyani shoppi?? (some paper) Maximum difference to Independence
  - goal is to find an axis in the space such that the projection of the stimulus on the axis creates a distribution that is maximally distribution form the distribution of all images. If everything is nicely gaussian, is the same as sta. Otherwise, it is dramatically different
  - reverse correlation techniques with white noise to get a RF, then use RF to simulate a neuron, don't work
  - however, natural images, with high dimensionality, are different from gaussian approached, so naïve approahces fail
  - RF of neuron with gaussian noise and reverse correlation gives wrong result; but using MID is better?

Adaption to natural stimulus statistics

- if the variability or contrast of input changes, then stimulus is bad, so you need to transform it to get features from the noise/signal. So now output of neuron is dependent on feature identification .
  - the neuron may only be able to represent a subset of the data; so most stimuli make neuron fire not at all, or at maximum firing rate. So you need to scale the stimulus of variability dynamically.
  - the neuron's range needs to match the range of input stimulus
  - If the range of the neuron is smaller than the range of the input, bad things happen at the edges
- The distribution that has maximum entropy, has maximum ability to describe x, is onstant, and doesn't have to be a boolean distribution.
  - applying the same idea to neural responses - maybe have a neuron that spikes at certain ranges, then count spike rate over time at time bins, and calculate the maximum entropy of the neuron, need to make sure all outputs as used equally
  - need to choose a input/output function in such a way thtat the responses of a neuron are all equal
    - from input, generate a histogram, then generate a CDF to get corresponding bins that need to be mapped to the output
    - therefore, all outputs of neuron are equally likely - histogram normalization
    - Is this what actual neurons are doing?

Do natural neurons use normalization to optimize neruon firing?

- Laughlin 1981
  - comparing contrast with response to input
  - find that flies optimally adapt
  - The optimal function is different on a bright sunny day vs a cloudy day
- Fairhalll 2001
  - showed randomly moving stimulus to fly, but random process is not homogenous, but variance of random process changed over time
  - scaled noise so that standard deviation of noise hanged over time, plotting input/output function results in getting neural response of neuron?? you
  - find that all input/output fiunctions overall all standard deviation bins agree
  - the nonlinearity of the fly changes such that the response of the neuron contains

maximal information about the stimulus

So mapping from input to spike is actually quite dynamic

Adaption to natural stimulus statistics

- in low signal to noise with low light, the optimal filter looks different from a high signal to noise with high light fitler
- what you find is that the neuron, when light is low…. averages over space to produce output.
- whereas the same neuron might use smoothing or sharpening depending on the signal/noise ratio
- filter changes, the input/output function doesn't change.
- to maximize information in image and output, you need to use a dynamic filter

Maximum entropy

- fixed range: constant distribution
- fixed variance: gaussian distribution
- If you have only positive values and you fix the mean: exponenital distribution
    - the distribution over the spike counts that maximizes the information the neuron can carry given fixed metabolic costs. For a neuron with an avg spike rate of x, what is P(r) to maximize information
    - simetimes the neuron spikes a lot, somtetimes it doesn't spike often, so on average its exponential
    - this is for a single neuron
- what about multiple neurons?
    - entropy of two variables \leq entropy of sum of variables separately
    - $E(x_1, x_2) \leq E(x_1) + E(x_2)$
    - redundacny reduction
        - the brain should learn recpetive fields and input output functions that reduce redundancy in neuron, de-correlate the neurons
        - the output of neurons depends on each other,
        - gives rise to indepened component analysis to explain arising of garbor shape

Hard to get enough data for full multi-dimension histograms

the histogram of neural responses, you need enough observations to estimate the distribution with a histogram,

Information methods don't scale well, intractable with more than a handful of neurons

Even if you hava lot of data, th eestimators are severely biased and easily give misleading results

- unbiased estimators don't really exist statistically

Independent Component Analysis is the future? (or at least the near future)

Computational Neuroscience

Decoding

- for next Wednesday, read Parker & Newsome 1998

How can we infer whats in thte outside world given the responses?

stimulues -> sensroy P(r|s) -> decision P(s|r) -> Decision

Moving dots task - subject shown randomly moving dots on a screen, some percentage are moving together in one direction (coherence)

- animal has to indicate what they see

Bayesian computation is computationally hard because R is a vector of all neural responses, with millions of neurons, incredibly high dimensionality, will never be able to fully characterize. So clearly the brain isn't doing this, so what do

- Factorization over neurons, assuming neurons are independent

- Population code, tuning curves,
  - on any one trial, you don't see the tuning curve, just the response of all neurons at any one time
  - labeling neurons with preffered directoin, plotting mean response of all neurons with the same stimulus. The neuron whose prefered direction is the stimulus will have the highest acivity at that stimulus??
  - Each of the neurons will spike with a certain number of spikes from a poisson distribution.....
  - so the simplest approach is to just look at the neuron which spikes the most at whichever stimulus of interest

Nature of the problem

- in repsonse to a stimulus with unknwn orientation s, you obseerve a pattern of activity r, what can you say about s given r?
- estimation theory - come up with a single value estimate s\hat from r
- bayesian approach - recover p(s|r), get a confidence estimate

Maximum likelihood,

- extract template/regression from population response
- then take template and find where it best matches the oberved data, overlay template onto observed data and get maximum likelihood estimate
- the maximum likelihood estimate is the value os s maximizing the likelihood p(r|s). Therefore we seek s\hat such that
  - $\hat{s}_{ML} = \arg\max P(r|s)$
  - $P(r|s) \approx poisson\ distribution\ = \prod_{all\ neurons} e^{\frac{(r_i - f(s))^2}{2\sigma^2}}$
  - $argmax\ = \sum -\frac{[r_i - f(s)]^2}{2\sigma^2} \approx MSE$
  - $\log \prod g_{i(x)} = \log \sum g_{i(x)}$
  - If the spiek rate is high, then can replace $\backslash sgima^2\ with\ r_i$

Doing n estimates gives n (slightly) different estimators

- is Estimator os s = s, then estimate is said to be unbiased
- if $sigma^2_{\hat{s}(s)}$ is as small as possible, is said to be ___ ??
- MSE = expected difference between estimate and truth
  - = bias + variability
  - = devaition on average over time + the varaince around that mean
- in general, the best estimator is bound by the Cramer-Rao bound
  - the lowest variance a given estimator can have in a given system
  - usually, estimator variance is larger than the CR variacnce

Fisher Information

- fundamentally different information from Shannon?? information
- defined as inverse of variance of efficient estiamtor (best estimator)
- is equal to some formula (see slides)
- equivalently = the expected value of the log of the likeihood of the second derivative of some ass probability
- for one single neuron, the Fisher information is given by (see slides)
    - $I_{i(s)} = \frac{f'2(s)^2}{f_{i(s)}} \approx d'^2$

??????? df

- I(S) = f(S)/\sigma^2
- the more neurons the better
- small variance is good
- large values are good