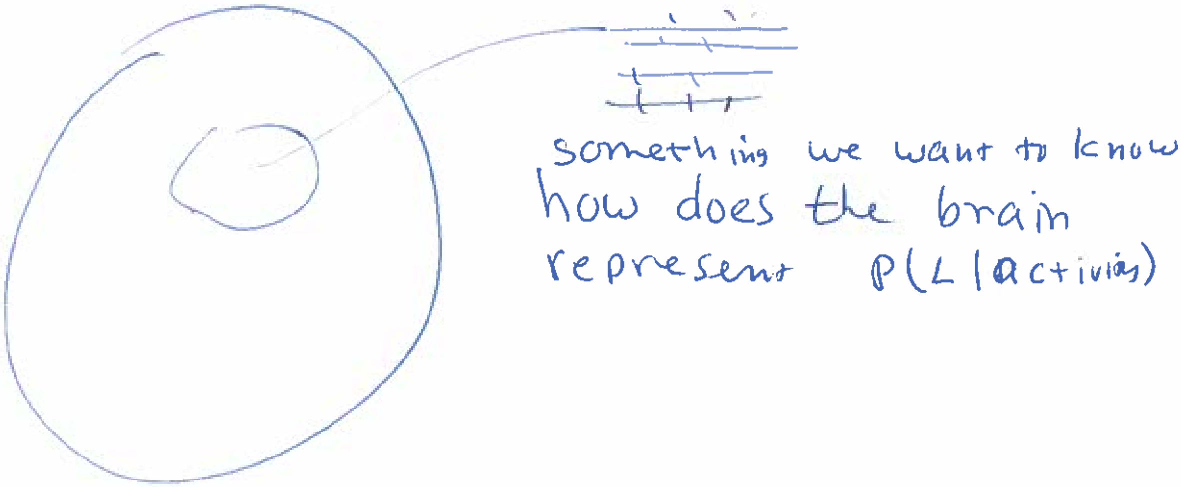
Oct. 26, 2017

Motivation



how does the brain represent $P(L \mid \text{activity})$

- too hard. Settle for $\hat{L}$ (an estimate of L) and its variance.
- knowing the variance – and how it evolves as the brain computes – is useful for understanding how computations can be robust to noise.

Example

RDK

- single neurons sometimes did as well as the monkey
- weak correlation between single neuron firing rate & choice (choice probability > 0.5)
- how do we understand this?

Refs:

Shadlen et al JN 16:1486–1510 (1996)

Shamir & Sompolinsky NC 18:1951-86 (2004)

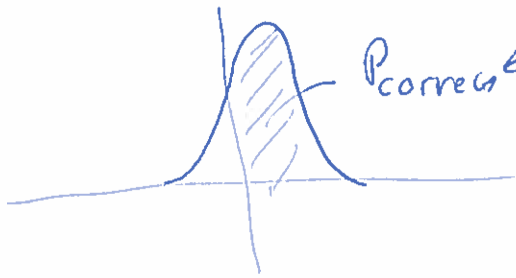
Ecker et al JN 31:14272-83 (2011)

Zylberberg et al PLoS CB 13: e1005497 (2017)

Tuning curve + noise

$$r_i = f_i(s) + \xi_i$$

$p(\underline{r}|s) \longrightarrow p(s|\underline{r}) \longrightarrow \boxed{\begin{matrix}\text{mean}\\ \text{variance}\end{matrix}} \longrightarrow$ choice



1. compute mean + variance vs. pop. size

monkey looks at whole population, so variance of single neuron can't be much lower than variance of whole population!!

<u>Gaussian assumption</u>

$$p(\underline{r}|s) \propto e^{-\frac{1}{2}\sum_{i=1}^{n} \frac{(r_i - f_i(s))}{\sigma_i^2}}$$

ML decoding $\hat{s}$ minimize

what's the variance?

$$p(\underline{r}|s) \sim e^{-\frac{1}{2}\sum_i \frac{\partial^2}{\partial s^2} \frac{(r_i - f_i(s))^2}{\sigma_i^2}\Big|_{s=\hat{s}} (s-\hat{s})^2}$$

$\underline{\underline{\propto n}}$

<u>as $n \to \infty$, animal almost perfect</u>

The problem is the independence assumption

$$r_i = f_i(s) + \xi_i \qquad \langle \xi_i \rangle = 0$$

$$\langle \xi_i \xi_j \rangle = \Sigma_{ij}$$

$$p(\underset{\sim}{r}|s) \propto e^{-\frac{1}{2}(\underset{\sim}{r} - \underline{f}(s)) \cdot \underline{\underline{\Sigma}}^{-1} (\underline{r} - \underline{f}(s))}$$

we'll return to this, but first, some intuition

$$\langle \xi_i \xi_j \rangle = \sigma^2 (1 + \rho (1 - \delta_{ij}))$$

Say $r_i = s + \xi_i$

linear tuning curves

$$\hat{s} = \frac{1}{n} \sum_{i=1}^{n} r_i = s + \frac{1}{n} \sum_i \xi_i$$

$$\mathrm{Var}[\hat{s}] = \frac{1}{n^2} \sum_{ij} \langle \xi_i \xi_j \rangle$$

$$= \frac{1}{n^2} \sum_i \langle \xi_i^2 \rangle + \frac{1}{n^2} \sum_{i \neq j} \langle \xi_i \xi_j \rangle$$

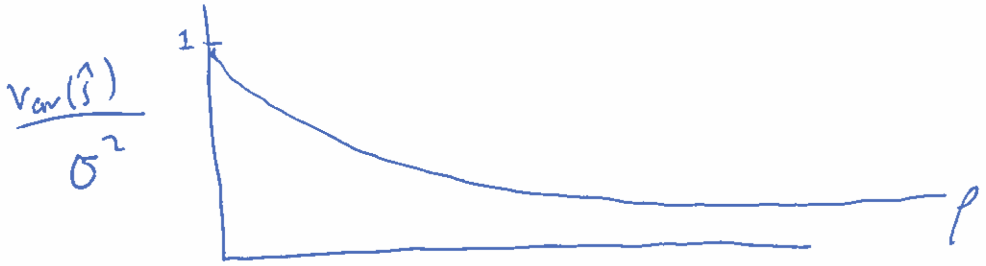
$$= \sigma^2 \left[\frac{1}{n} + \frac{n(n-1)}{n^2} \rho \right]$$

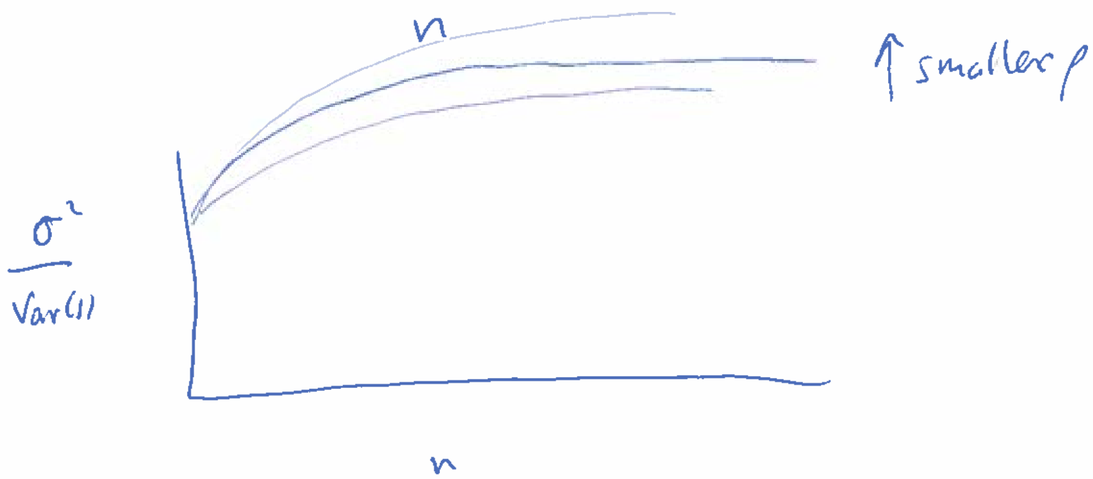
$$= \sigma^2 \left[\frac{1-\rho}{n} + \rho \right]$$

important when $\rho \sim \frac{1}{n}$, $n \sim \frac{1}{\rho}$

usual $\frac{1}{n}$ scaling for ind. neurons

even small correlations matter





- correlations also explain choice probabilities, at least in principle: higher r_i on one neuron $\Rightarrow$ higher r_i on lots of neurons

- to discuss arbitrary correlational structures, we need a better set of tools.

- Fisher info !!!

$$p(\underline{r}|s) = e^{-\psi(\underline{r}, s)}$$

$$\hat{s} : \left.\frac{\partial \psi}{\partial s}\right|_{s=\hat{s}} = 0$$

$$p(\underline{r}|s) = e^{-\psi(r,\hat{s}) - \frac{1}{2}\left.\frac{\partial^2 \psi}{\partial \hat{s}^2}\right|_{s=\hat{s}} (s-\hat{s})^2}$$

$$\mathrm{Var}^{(\hat{s},\underline{r})} \approx \frac{1}{\left.\frac{\partial^2 \psi(s,\underline{r})}{\partial s^2}\right|_{s=\hat{s}}} = \frac{1}{-\frac{\partial^2 \log p(\underline{r}|\hat{s})}{\partial \hat{s}}}$$

would like a quantity that tells us ~~how~~ about the variance on aver.

$$\overline{\mathrm{Var}(\hat{s})} = \frac{1}{-\left\langle \frac{\partial^2 \log p(\underline{r}|\hat{s})}{\partial \hat{s}^2} \right\rangle}$$

$I(s)$ = Fisher info

$\hat{s}(\underline{r})$

$$\mu(s) = \int d\underline{r}\ \hat{s}(\underline{r})\ p(\underline{r}|s)$$

$$0 = \int d\underline{r}\ \left(\hat{s}(\underline{r}) - \mu(s)\right) p(\underline{r}|s)$$

$$0 = \frac{\partial}{\partial s} \int d\underline{r}\ \left(-\mu'(s)\right) p(\underline{r}|s) + \left(\hat{s}(r) - \mu(s)\right) p(r|s)\, \partial \log p$$

$$\mu(s) = s + b(s)$$

$$1 + b'(s) = \langle (\hat{s} - \mu)\, \partial_s \log p(r|s) \rangle$$

$$(1+b')^2 \leq \langle (\hat{s}-\mu)^2 \rangle \langle (\partial_s \log p)^2 \rangle$$

$$\mathrm{Var} \geq \frac{(1+b'(s))^2}{\langle (\partial_s \log p)^2 \rangle}$$

$$-\langle \partial_s^2 \log p \rangle = I$$

multi-variable anal

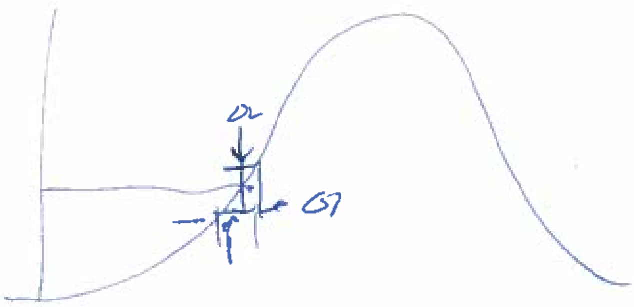
$$I_{ij} = -\left\langle \frac{\partial^2 \log p(\underline{r}|\underline{s})}{\partial s_i \partial s_j} \right\rangle$$

ind. Gaussian

$$e^{-\frac{1}{2}\sum_i \frac{(r_i - f_i(s))^2}{\sigma_i^2}}$$

$$\left\langle \frac{\partial^2}{\partial s^2} \right\rangle = \cancel{+\sum \frac{(r_i - f_i(s)) f_i''}{\sigma_i^2}} - \sum_i \frac{f_i'^2}{\sigma_i^2} \quad dn$$

intuition:



$$\frac{\Delta r}{\Delta s} = \frac{df}{ds} \qquad \mathrm{var}(s) = \frac{\mathrm{var}(r)}{f'^2}$$

$$e^{-\frac{1}{2}\frac{(s-\hat{s})^2 f'(s)^2}{\sigma^2}}$$

multiple sources

$$\sum_i \qquad \hat{s}_i \quad \frac{f'^2}{\sigma^2}$$

$$e^{-\sum_i \frac{(s-\hat{s}_i)^2 I_i}{}}$$

ML estimate

$$\hat{s} = \frac{\sum_i \hat{s}_i I_i}{\sum_i I_i} \qquad \frac{1}{\sigma^2} = \sum I_i$$

Poisson $P(r|s) = \frac{f(s)^r e^{-f(s)}}{r!}$

$$L = r \log f - f$$

$$\frac{\partial L}{\partial s} = \frac{r f'}{f} - f'$$

$$\frac{\partial^2 L}{\partial s^2} = \frac{r f''}{f} - \frac{r f'^2}{f^2} - f''$$

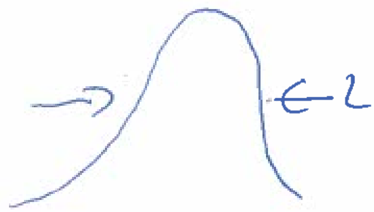
$$-\langle\ \rangle = \frac{f'^2}{f} \quad \frac{f''}{f}$$

Linear Fisher

1-D: steeper is better!!

high D: not true

$$e^{-\sum_i \frac{(\underline{r} - \underline{f}(\underline{s}))^2}{2\sigma^2}}$$



$$I = \sum_i \frac{f'(s)^2}{2\sigma^2}$$

$$(f')^2 \sim \frac{1}{L^2}$$

neurons in s ~ L^D

$$I \sim L^{D-2}$$

$$e^{-\frac{1}{2}(\underline{r}-\underline{f})\cdot\underline{\underline{\Sigma}}^{-1}(\underline{r}-\underline{f})}$$

$$I = \underline{f}'\cdot\underline{\underline{\Sigma}}^{-1}\cdot\underline{f}'$$

$$\underline{\underline{\Sigma}}^{-1} = \sum_k \lambda_k^{-1}\underline{v}_k\underline{v}_k$$

perh alm

$$I = \sum_k \frac{(\underline{f}'\cdot\underline{v}_k)^2}{\lambda_k} = |f'|^2 \sum_k \frac{\cos^2\theta_k}{\lambda_k}$$

Shadlen:

$$\frac{\underline{\underline{\Sigma}}}{\sigma^2} = \begin{bmatrix} 1 & & \rho \\ & \ddots & \\ \rho & & 1 \end{bmatrix} = (1-\rho)I + \begin{bmatrix} \rho & \rho & \rho \\ \rho & \rho & - \end{bmatrix}$$

$$= (1-\rho)I + \rho\,\underline{1}\,\underline{1}$$

$$\Sigma\cdot\underline{1} = (N\rho + (1-\rho))\underline{1}$$

$$\Sigma\cdot\underline{1}^\perp = (1-\rho)\underline{1}^\perp$$

$$\lambda_0 = N\rho + (1-\rho)$$

$$\lambda_1 = (1-\rho)$$

$$\underline{f}' = \underline{1} \qquad \cos^2\theta_0 = 1$$

$$\cos^2\theta_{k>1} = 0$$

S & S Eckers (find natl : $\rho(\cos\theta, \theta\,\%) \neq 0$

$\Rightarrow$ generically $I \propto n$!!

a problem: $n \to \infty$, I must saw

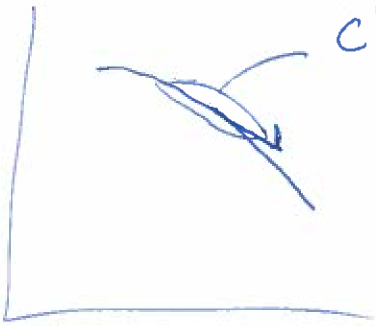
$$I = |f'|^2 \sum_k \frac{\cos^2\theta_k}{\lambda_k}$$

$\hookrightarrow O(n)$ $k=0$

$O(1/n)$ $k \neq 0$

- different correlation!!
- guarantee by OPI
- what do the look ?

$V_\infty \sim f'$



c's have to line up

~~a problem~~

The good news:

nearly noise-free computations

$$x = f(s) + \xi_x$$

$$y = w \cdot x + \xi_y$$

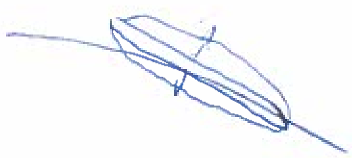
$$= w \cdot f(s) + w \cdot \xi_x + \xi_y$$

$$I = f' w^T \cdot [w \cdot \Sigma_x \cdot w + \Sigma_y]^{-1} \cdot w \cdot f'$$

$$= f' [\Sigma_x + (w^T \cdot \Sigma_y^{-1} \cdot w)^{-1}] \cdot f'$$

$O(n)$ $\quad$ $O(1)$ multiplier rech $\quad$ added in

$$\Sigma_s \quad \xi_x + \xi_y' \quad \sum_k \frac{\cos^2\theta}{\lambda_k + \sigma^2}$$



the bad news: can't encode probabilities!!

h: $\Sigma_{ii} = \frac{1}{J_{ii} - J_{ij} \tilde{J}_{jk}^{-1} J_{ki}}$