Coding with Stimulus Dependent Covariance

Let us consider an interesting variant of the classic population coding problem for N neurons. Let the firing rate of neuron i be r_i, and given a stimulus θ, suppose the response r_i is distributed according to a Gaussian distribution with a θ-dependent covariance matrix:

$$\mathrm{p}(r|\theta) = \frac{1}{(2\pi)^{N/2}\sqrt{\det(C(\theta))}} \exp\left[-\frac{1}{2}(r_i - \mu_i)C^{-1}_{ij}(r_j - \mu_j)\right].$$

For simplicity, we take $\mu_i = \mu$. And as always, assume $N \gg 1$.

Let's begin by finding the Fisher information, $J(\theta)$.

(a) Prove the matrix identity

$$\frac{\partial A^{-1}}{\partial\theta} = -A^{-1}\frac{\partial A}{\partial\theta}A^{-1}.$$

(b) Prove the matrix identity

$$\frac{\partial \det(A)}{\partial\theta} = \det(A)\mathrm{tr}\left(A^{-1}\frac{\partial A}{\partial\theta}\right).$$

(c) Use these identities to show that

$$J(\theta) = \frac{1}{2}\mathrm{tr}\left(\left(C^{-1}\frac{\partial C}{\partial\theta}\right)^2\right).$$

Now, we wish to understand the consequences of the form of $J(\theta)$ above. For simplicity, let us write

$$C_{ij}(\theta) = \delta_{ij}D(\theta - \phi_i) + S_{ij}$$

where

$$S_{ij} \equiv S(\phi_i - \phi_j).$$

and S is a smooth function with a maximum at 0. As usual, we are letting

$$\phi_j \equiv \frac{2\pi \mathrm{i} j}{N}$$

To find a lower bound on $J(\theta)$ when we have no hope of diagonalizing the covariance matrix, we need to find an analogue of the "population vector" that is used in the case of the stimulus-dependent rate. A suitable choice will be

$$R_i = \left(\sum_{j=1}^{N} W_{ij} r_j\right)^2$$

with

$$W_{ij} = \frac{1}{N}\sum_{|n|>p} \mathrm{e}^{\mathrm{i}n(\phi_i - \phi_j)}$$

for some value p. Then, we define

$$z = \sum_{i=1}^{N} R_i \mathrm{e}^{\mathrm{i}\phi_i}.$$

This will be our population vector.

(d) Suppose that $N \gg p$, but p is still "large." Show that

$$\langle z \rangle \approx N \mathrm{e}^{\mathrm{i}\theta} \int_0^{2\pi} \frac{\mathrm{d}\phi}{2\pi} \mathrm{e}^{\mathrm{i}\phi} D(\phi)$$

if we choose a "large" enough value of p – explain in what sense this approximation will be true.

(e) Argue that $\mathrm{Var}(z) \sim N$. You do not need to focus on the coefficient of proportionality anymore.

(f) Conclude that the Fisher information $J(\theta) \sim N$ by combining the previous parts.

(g) The argument above breaks down for $N \sim N_{\mathrm{sat}}(p)$, a p-dependent saturation limit. You do not need to justify this statement – trust me that it's true! The reason that this happens is that the argument made in the first part of this section breaks down – explain why! Do you think this argument means that $J(\theta)$ asymptotes to a constant for large N – why or why not?

Recall that in the case of a θ-independent C, but θ-dependent r, the Fisher information was only $\mathrm{O}(1)$. This suggests that it is actually far more preferable for population coding to actually encode information about the stimulus in the covariance matrix.