Assignment 6
Theoretical Neuroscience

Peter Dayan, Maneesh Sahani, Misha Ahrens (ahrens@gatsby.ucl.ac.uk)

November 11, 2005

1. **Fisher Information**

   Prove the two results from lecture:

   (a)
   $$\left\langle \frac{d}{d\theta} \log P(y|\theta)\bigg|_{\theta^*} \right\rangle_{y\sim P(y|\theta^*)} = 0$$

   (b)
   $$\left\langle \left(\frac{d}{d\theta} \log P(y|\theta)\bigg|_{\theta^*} \right)^2 \right\rangle_{y\sim P(y|\theta^*)} = -\left\langle \frac{d^2}{d\theta^2} \log P(y|\theta)\bigg|_{\theta^*} \right\rangle_{y\sim P(y|\theta^*)}$$

2. **Population Coding**

   Shadlen and collaborators have claimed that if the activities of neurons in population codes are corrupted by *correlated* noise, then there is a sharp limit to the useful number of neurons in the population. *Prima facie* this is wrong – the stronger the correlations, the lower the entropy of the noise, and therefore the stronger the signal.

   Resolve this issue for the case of additive and multiplicative noise by considering the following three models for the noisy activities r_1 and r_2 of two neurons which form a population code for a real-valued quantity x:

   $$\text{a)} \quad \begin{cases} r_1^a = x + \epsilon_1 \\ r_2^a = x + \epsilon_2 \end{cases} \tag{1}$$

   $$\text{b)} \quad \begin{cases} r_1^b = x(1-\delta) + \epsilon_1 \\ r_2^b = x(1+\delta) + \epsilon_2 \end{cases} \tag{2}$$

   $$\text{c)} \quad \begin{cases} r_1^c = x(1-\delta)(1+\eta_1) \\ r_2^c = x(1+\delta)(1+\eta_2) \end{cases} \tag{3}$$

   where $\delta \neq 0$ is known, and, ϵ and η are Gaussian, with mean 0 and covariance matrices:

   $$\Sigma = \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix}$$

   (a) What is the maximum likelihood (ML) estimator for x on the basis of r_1 and r_2 in each case?

   (b) What is the appropriate measure of the expected accuracy of the ML estimate, and why?

(c) How does the expected accuracy in each case depend on the degree of correlation c?

(d) What conclusions would you draw about the clash between Shadlen and common sense?

3. **Reinforcement learning.** *Please note that the equation numbers refer to equations in Dayan and Abbott, chapter 9.*

   Consider the case of partial reinforcement (studied in figure 9.1) in which reward $r = 1$ is provided randomly with probability p on any given trial. Assume that there is a single stimulus with $u = 1$, so that $\epsilon\delta u$, with $\delta = r - v = r - wu$, is equal to $\epsilon(r - w)$. By considering the expected value $\langle w + \epsilon(r - w)\rangle$ and the expected square value $\langle (w + \epsilon(r - w))^2\rangle$ of the new weights, calculate the self-consistent equilibrium values of the mean and variance of the weight w. What happens to your expression for the variance if $\epsilon = 2$ or $\epsilon > 2$? To what features of the learning rule do these effects correspond?

4. Implement a stochastic three-armed bandit using the indirect actor and the action choice softmax rule 9.12. Let arm a produce a reward of p_a, with $p_1 = 1/4, p_2 = 1/2, p_3 = 3/4$, and use a learning rate of $\epsilon = 0.01, 0.1, 0.5$ and $\beta = 1, 10, 100$. Consider what happens if after every 250 trials, the arms swap their reward probabilities at random. Averaging over a long run, explore to see which values of ϵ and β lead to the greatest cumulative reward. Can you account for this behavior?

5. Repeat the above excercise using the direct actor (with learning rule 9.22). For $\bar{r}$, use a low-pass filtered version of the actual reward, which is obtained by using the update rule

   $$\bar{r} \rightarrow \lambda\bar{r} + (1 - \lambda)r$$

   with $\lambda = 0.95$. Study the effect of the different values of ϵ and β in controlling the average rate of rewards when the arms swap their reward probabilities at random every 250 trials.