Covariance and Correlation: Assume that we have recorded two neurons in the two-alternative-forced choice task discussed in class. We denote the firing rate of neuron 1 in trial $i$ as $r_{1,i}$ and the firing rate of neuron 2 as $r_{2,i}$. We furthermore denote the averaged firing rate of neuron 1 as $\bar{r}_1$ and of neuron 2 as $\bar{r}_2$. Let us “center” the data by defining two data vectors

$$x = (r_{1,1} - \bar{r}_1, r_{1,2} - \bar{r}_1, r_{1,3} - \bar{r}_1, \ldots, r_{1,N} - \bar{r}_1)$$

$$y = (r_{2,1} - \bar{r}_2, r_{2,2} - \bar{r}_2, r_{2,3} - \bar{r}_2, \ldots, r_{2,N} - \bar{r}_2)$$

(a) Show that the variance of the firing rates of the first neuron is $\text{Var}(r_1) = \frac{1}{N}||x||^2$.

(b) Compute the cosine of the angle between $x$ and $y$. What do you get?

(c) What are the maximum and minimum values that the correlation coefficient between $r_1$ and $r_2$ can take? Why?

(d) What do you think the term “centered” refers to?

Bayes theorem: The theorem of Bayes summarizes all the knowledge we have about the stimulus by observing the responses of a set of neurons, independent of a specific decoding rule. To get a better intuition about this theorem, we will look at the motion discrimination task again, and compute the probability that the stimulus moved to the left (+) or right (-). For a stimulus $s = \{+, -\}$, and a firing rate response $r$ of the neuron, the theorem of Bayes reads

$$p(s|r) = \frac{p(r|s)p(s)}{p(r)}.$$

Here, $p(r|s)$ is the probability that the firing rate is $r$ if the stimulus was $s$. The respective distribution can be measured and we assume that it follows a Gaussian probability density with mean $\mu_s$ and standard deviation $\sigma$,

$$p(r|s) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r - \mu_s)^2}{2\sigma^2}\right).$$

The relative frequency with which the stimuli (leftward or rightward motion, + or −) appear is denoted by $p(s)$, often called the prior probability or, for short, the prior. The distribution $p(r)$ denotes the probability of observing a response $r$, independent of any knowledge about the stimulus.

(a) How can you calculate $p(r)$?

(b) The distribution $p(s|r)$ is often called the posterior probability or, for short, the posterior. Calculate the posterior for the stimulus $s = +$. Sketch the posterior $p(s = +|r)$, or, for short, $p(+|r)$ as a function of $r$, assuming a prior $p(+) = p(-) = 1/2$. Draw the posterior $p(-|r)$ into the same plot.
(c) What happens if you change the prior? Investigate how the posterior changes if \( p(+) \) becomes much larger than \( p(-) \) and vice versa. Make a sketch similar to (c).

(d) Let us assume that we decide for leftward motion whenever \( r > \frac{1}{2}(\mu_+ + \mu_-) \). Interpret this decision rule in the above plots. How well does this rule do depending on the prior? What do you lose when you move from the full posterior to a simple decision (=decoding) rule?

(3) **Linear discriminant analysis (advanced):** Let us redo the calculations for the case of \( N \) neurons. If we write \( r \) for the vector of firing rates observed, then Bayes theorem now reads:

\[
p(s|r) = \frac{p(r|s)p(s)}{p(r)}.
\]

We assume that the distribution of firing rates again follows a Gaussian so that

\[
p(r|s) = \frac{1}{(2\pi)^{N/2} \sqrt{\det C}} \exp \left( -\frac{1}{2} (r - m_s)^T C^{-1} (r - m_s) \right)
\]

where \( m_s \) denotes the mean of the density for stimulus \( s \), and \( C \) is the covariance matrix.

(a) Compute the log-likelihood ratio

\[
l(r) = \log \frac{p(+|r)}{p(-|r)}.
\]

(b) Assume that \( l(r) = 0 \) is the decision boundary, so that any firing rate vector \( r \) giving a log-likelihood ratio larger than zero is classified as coming from the stimulus +. Compute a formula for the decision boundary. What shape does this boundary have?

(c) Assume that \( p(+) = p(-) = 1/2 \). Assume we are looking only at two neurons that are uncorrelated so that the covariance matrix is

\[
C = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}
\]

Sketch the decision boundary for this case.