MOD202 - Exercise 4

Ciprian Bangu

March 25 2024

1 The Nightly War of Moths and Bats

a) The intensities of the bat's sonar on the moth's ears are given by the following equations:

For the moth's right ear:

$$I_R = \frac{I}{r}(\cos^2(\varphi/2) + \alpha) \tag{1}$$

For the moth's left ear:

$$I_L = \frac{I}{r}(\sin^2(\varphi/2) + \alpha) \tag{2}$$

where r is the distance between the bat and the moth, I is the intensity of the bat's sonar, φ is the angle between the bat's sonar and the moth's ears, and α is a constant.

Figure 1 shows the intensity of the bat's sonar on the moth's ears as a function of φ, where the total intensity of the bat's sonar $I = 1$, $\varphi \in [0, 180]$, r is fixed at $1m$, and $\alpha = 0$.

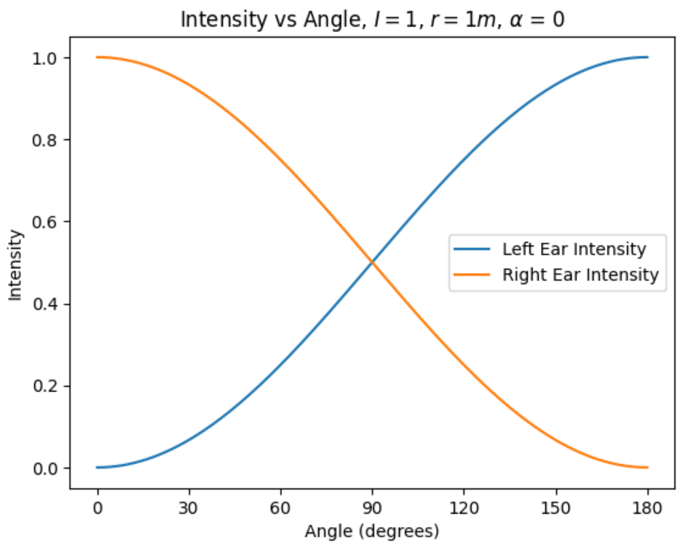

Figure 1: The intensity of the bat's sonar on the moth's ears as a function of φ

We see that when the sound comes from the right, i.e., $\varphi < 90°$ the intensity on the right ear is higher than the intensity on the left ear, and vice versa. When the sound comes from directly in front of the moth, i.e., at an angle of $90°$, the intensities on both ears are equal.

b) Given a fixed φ and a varying distance r, the intensity of the bat's sonar on the moth's ears is shown in Figure 2.

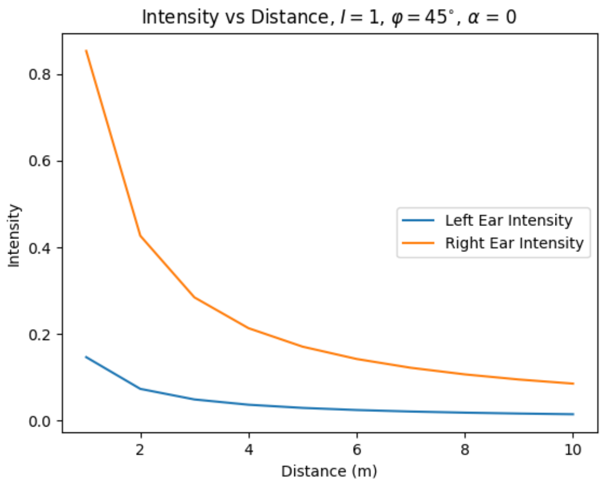

Figure 2: Intensity of the bat's sonar on the moth's ears as a function of r

We see that, given the angle is fixed, and less than $90°$, the intensity of the bat's sonar recieved on the right ear is always higher than the intensity recieved on the left ear. This corrsponds to the shape of the curves in Figure 2, which shows that $I_L < I_R$ for $\varphi < 90°$. The opposite would hold for a fixed angle greater than $90°$. If the angle is $90°$, the intensities on both ears as a funciton of the distance are always equal.

Moreover, we see that the intensity of the bat's sonar on the moth's ears decreases as the distance between the bat and the moth increases. This is because the intensity of the bat's sonar is inversely proportional to the distance between the bat and the moth, as shown in equations (1) and (2).

2 Parameter Estimation

a) Given that the Mean Square Error $E(w)$ is defined as:

$$E(w) = \sum_{i=1}^{M} (r_i - ws_i)^2 \tag{3}$$

it's derivative with respect to w is:

$$\begin{aligned}
\frac{dE(w)}{dw} &= \frac{d}{dw} \sum_{i=1}^{M} (r_i - ws_i)^2 \\
&= \sum_{i=1}^{M} \frac{d}{dw} (r_i - ws_i)^2 \\
&= \sum_{i=1}^{M} \frac{d}{du} (u)^2 \frac{du}{dw} && \text{Substituting } r_i - ws_i = u \\
&= \sum_{i=1}^{M} 2u \frac{du}{dw} \\
&= \sum_{i=1}^{M} 2(r_i - ws_i)(-s_i) && \text{differentiating with respect to } w
\end{aligned}$$

Thus, the derivative of $E(w)$ with respect to w is:

$$\frac{dE(w)}{dw} = -2 \sum_{i=1}^{M} (s_i)(r_i - ws_i) \tag{4}$$

b) The funciton $E(w)$ and its derivative $\frac{dE(w)}{dw}$ are shown in Figure 3.

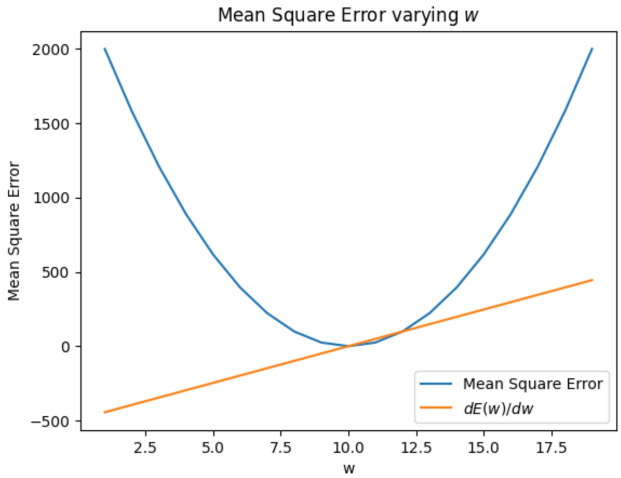

Figure 3: The Mean Square Error and it's derivative.

c) We can evaluate a guess by Mr. Smarty Pants by inspecting the derivative of the Mean Square Error at his guess for the weight, w_{sp}. The value he is trying to guess, w_{opt}, is the value of w that minimizes the Mean Square Error, i.e., w such that $E(w) = E(w)_{min}$. From Fig. 3, we can see that when the MSE is at its minimum, the derivative of the MSE is equal to 0. Moreover, as it approaches its minimum from the left, it's derivative is negative; as it approaches its minimum from the right, it's derivative is positive. Therefore, if $\frac{dE(w)}{dw}(w_{sp}) = 0$, Mr. Smarty Pants' guess is correct and $w_{sp} = w_{opt}$. However, If $\frac{dE(w)}{dw}(w_{sp}) < 0$, then his guess is too small. Conversely, if $\frac{dE(w)}{dw}(w_{sp}) > 0$, then his guess is too large.

d) The above analysis leads to a rule for updating the guess in order to minimize $E(w)$. Namely, we know that if the $\frac{dE(w)}{dw}$ for some w is greater than 0, we need to decrease w in order to minimize $E(w)$. Conversely, if $\frac{dE(w)}{dw}$ is less than 0, we need to increase w in order to minimize $E(w)$. And, if $\frac{dE(w)}{dw} = 0$, then we have found the minimum of $E(w)$. Therefore, we can update our guess of w based on the value of $\frac{dE(w)}{dw}$. Specifically, by subtracting $\frac{dE(w)}{dw}$ from the value of $E(w)$ for our current guess:

$$w_{new} = w_{old} - \frac{dE(w)}{dw}$$

When $\frac{dE(w)}{dw}$ is positive, we know the w chosen is too large, and thus the update rule will provide a smaller value of w. Conversely, when $\frac{dE(w)}{dw}$ is negative, we know the w chosen is too small, and thus the update rule will provide a larger value of w.

However, the above rule will not work in its current form, as $\frac{dE(w)}{dw}$ will tend to be large for bad guesses of w. Therefore, we need to add a scaling factor η to manage the size of the update, and diminsh the impact of bad guesses. The update rule then becomes:

$$w_{new} = w_{old} - \eta \frac{dE(w)}{dw} \tag{5}$$

where η is a small positive number, e.g., $\eta = 0.01$.

Thus, when we iterate this update equation, $\frac{dE(w)}{dw}$ will eventually converge to 0. And, since the derivative of the Mean Square Error is 0 at the minimum, the value of w that minimizes the Mean Square Error will be found.

This is similar to the Perceptron learning rule, where the weight vetor is updated based on the error in the output.