# 1 Earthquakes

## 1.1 Earthquakes 1

In this question there are two explosions at locations $s_1$ and $s_2$ and the observed value at sensor $i$ is

$$v_i = \frac{1}{d_i^2(1) + 0.1} + \frac{1}{d_i^2(2) + 0.1} + \sigma\epsilon_i$$

where $d_i(1)$, $d_i(2)$ is the distance from explosion 1, 2 to the sensor $i$ respectively; $\sigma$ is the standard deviation of the Gaussian sensor noise and the noise $\epsilon_i$ is drawn from a zero mean unit variance Gaussian independently for each sensor. We assume that the prior locations of the explosions are independent and uniform.

**1.**

In this question we are asked to calculate the posterior $p(s_1|\mathbf{v})$.
First of all, let us find the distribution of $\mathbf{v}$.
We know that the signal is measured with Gaussian noise with a standard deviation of $\sigma$ and the noise $\varepsilon_i$ is drawn from a zero mean unit variance Gaussian independently for each sensor. This means that the observed value $v_i$ at sensor of index $i$ follows a Gaussian distribution.

$$\varepsilon_i \sim N(0, 1)$$

$$\therefore \quad E[v_i] = E\left[\frac{1}{d_i^2(1) + 0.1} + \frac{1}{d_i^2(2) + 0.1} + \sigma\varepsilon_i\right] = \frac{1}{d_i^2(1) + 0.1} + \frac{1}{d_i^2(2) + 0.1}$$

and,

$$V[v_i] = V\left[\frac{1}{d_i^2(1) + 0.1} + \frac{1}{d_i^2(2) + 0.1} + \sigma\varepsilon_i\right] = \sigma^2$$

$$\therefore \quad v_i \sim N\left(\frac{1}{d_i^2(1) + 0.1} + \frac{1}{d_i^2(2) + 0.1}, \sigma^2\right)$$

In order to calculate the posterior, we use Bayes theorem :

$$p(\theta \mid x) = \frac{f(x \mid \theta)\, p(\theta)}{\sum_\theta (f(x \mid \theta) p(\theta))}$$

where, $p(\theta \mid x)$ is the posterior distribution of $\theta$ given the data $x$. It represents the uncertainty about the parameter $\theta$ given the data x. $p(\theta)$ is the prior distribution. It represents the uncertainty about the true value of $\theta$ before collecting the data. $f(x \mid \theta)$ is the likelihood and $\sum(f(x \mid \theta)p(\theta))$ is the marginal likelihood. In practice we use $p(\theta \mid x) \propto f(x \mid \theta)\, p(\theta)$ .

Here,

$$p(S_1 \mid \mathbf{v}) \propto \sum_{S_2} p(\mathbf{v}, S_1, S_2)$$

Assuming that the prior locations of the explosions are independent and uniform (according to the spiral coordinate system):

$$\therefore p(S_1 \mid \mathbf{v}) \propto \sum_{S_2} \prod_{i=1}^{N} p(v_i, S_1, S_2) \quad , \ i = 1, ...N \ , \ N = 30$$

Then we can just normalise to compute the posterior:

$$\therefore p(S_1 \mid \mathbf{v}) = \frac{\sum_{S_2} \prod_{i=1}^{N} p(v_i \mid S_1, S_2)}{\sum_{S_1,S_2} \prod_{i=1}^{N} p(v_i \mid S_1, S_2)}$$

To use this formula we need to have a discrete number of possible locations S1 and S2 that approximate the continuous case (therefore a large number of possible locations). We therefore considered polar coordinates with a spiral which contains 2000 evenly spaced possible earthquake locations within the earth. This also provides us with a 1D case, therefore giving only one coordinate to each location. The marginal probability of the location S is constant and uniform across the spiral.

The 2000-dimensional vector output $p(s_1|\mathbf{v})$ can be returned using the code in the Appendix. The visualisation of this posterior is plotted in Figure 1.

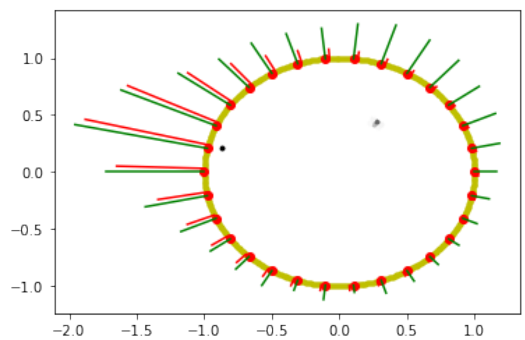


Figure 1: Plot of the maximum posterior $p(S_1 \mid \mathbf{v})$ for the scenario of two earthquakes (darker corresponds to a higher probability). This is based on a spiral coordinate system with $\sigma = 0.2$. The red points are the locations of the surface sensors. The observed (noisy) measurements at each sensor are represented by the green lines, and the true (unknown) blast values are denoted by the red lines.

**2.**

In this question we are asked to report the value of $(log(p(v|H_2)) - log(p(v|H_1))$.

Notice that, for the set $H_2$, $v_i \sim N(\frac{1}{d_i^2(1)+0.1} + \frac{1}{d_i^2(2)+0.1}, \sigma^2)$ whilst, for the set $H_1$, $v_i \sim N(\frac{1}{d_i^2(1)+0.1}, \sigma^2)$
Thus,

$$p(\mathbf{v} \mid H_2) = \sum_{S_1, S_2} p(\mathbf{v}, S_1, S_2 \mid H_2)$$

Assuming independence between S1 and S2:

$$\therefore p(\mathbf{v} \mid H_2) = \sum_{S_1, S_2} p(\mathbf{v} \mid S_1, S_2, H_2) p(S_1) p(S_2)$$

Likewise,

$$p(\mathbf{v} \mid H_1) = \sum_{S_1} p(\mathbf{v}, S_1 \mid H_1)$$

$$\therefore p(\mathbf{v} \mid H_1) = \sum_{S_1} p(\mathbf{v} \mid S_1, H_1) p(S_1)$$

Notice that : $p(S_1) = p(S_2) = \frac{1}{2000}$

We use the model used in subquestion 1 to compute $p(\mathbf{v} \mid H_2)$ and create a new model based on the same spiral polar coordinate idea but with only one earthquake location S1 to compute $p(\mathbf{v} \mid H_1)$. The code can be found in the Appendix and outputs **infinity**. It is worth noting that this is due to the probability of $p(\mathbf{v}|H_1)$ being so small that Python approximates its value to 0 therefore returning infinity for $log\left(\frac{p(\mathbf{v}|H_2)}{p(\mathbf{v}|H_1)}\right)$ since it is the denominator of a log ratio.

**3.**

$p(\mathbf{v} \mid H_2)$ and $p(\mathbf{v} \mid H_1)$ both compute the probability that the given observed values have indeed been generated from either one (H1) or two explosions (H2). Since prior probabilities are equal,

$$\frac{p(H_2|\mathbf{v})}{p(H_1|\mathbf{v})} = \frac{p(\mathbf{v}|H_2)p(H_2)}{p(\mathbf{v}|H_1)p(H_1)} = \frac{p(\mathbf{v}|H_2)}{p(\mathbf{v}|H_1)}$$

Also,

$$log(p(\mathbf{v}|H_2)) - log(p(\mathbf{v}|H_1)) = log\left(\frac{p(\mathbf{v}|H_2)}{p(\mathbf{v}|H_1)}\right)$$

Therefore, it computes the log of the ratio of having two explosions compared to one given our observed data. In our case, given our observed data, the output is a positive answer therefore showing that the probability of two explosions was higher than the probability of one. More than positive it actually outputs infinity as we discussed earlier due to the Python approximation to 0 for $p(\mathbf{v}|H_1)$ therefore showing that the model clearly predicts a higher probability for two explosions in comparison with only one.

**4.**

As we saw in subquestion 2, the computation is done by marginalising the likelihood over all the locations $S_1, ..., S_K$:

$$log(p(v|H_K)) = log\left(\sum_{S_{1:K}} \prod_{i=1}^{N} p(v_i \mid S_{1:K})\right)$$

Therefore, for K explosions, the computation will create a loop iterating for each possible location here of size n=2000 indented for each explosion $S_{1:K}$. Therefore the computational complexity will be of order $n^K$.

## 1.2 Earthquakes 2

This question is similar to the question in 2.1, except a different kind of explosion sensor measures the mean of two incoming explosions.

$$v_i = \frac{0.5}{d_i^2(1) + 0.1} + \frac{0.5}{d_i^2(2) + 0.1} + \sigma\epsilon_i$$

We evaluate the probabilities using the mean observed value at sensor i (instead of the regular observed value).

**1.**

In this question, we are asked to calculate the posterior $p(s_1|\mathbf{v})$ given : $v_i \sim N\left(\frac{0.5}{d_i^2(1)+0.1} + \frac{0.5}{d_i^2(2)+0.1}, \sigma^2\right)$. We proceed in a similar way as we did for question 2.1.1

In order to calculate the posterior, we use Bayes theorem :

$$p(\theta \mid x) = \frac{f(x \mid \theta)\, p(\theta)}{\sum_\theta (f(x \mid \theta) p(\theta))}$$

where, $p(\theta \mid x)$ is the posterior distribution of $\theta$ given the data $x$. It represents the uncertainty about the parameter $\theta$ given the data x. $p(\theta)$ is the prior distribution. It represents the uncertainty about the true value of $\theta$ before collecting the data. $f(x \mid \theta)$ is the likelihood and $\sum(f(x \mid \theta)p(\theta))$ is the marginal likelihood. In practise we use $p(\theta \mid x) \propto f(x \mid \theta)\, p(\theta)$ .

$$p(S_1 \mid \mathbf{v}) \propto \sum_{S_2} p(\mathbf{v}, S_1, S_2)$$

Assuming that the prior locations of the explosions are independent and uniform (according to the spiral coordinate system):

$$\therefore p(S_1 \mid \mathbf{v}) \propto \sum_{S_2} \prod_{i=1}^{N} p(v_i, S_1, S_2) \quad , \; i = 1, ...N \; , \; N = 30$$

Then we can just normalise to compute the posterior:

$$\therefore p(S_1 \mid \mathbf{v}) = \frac{\sum_{S_2} \prod_{i=1}^{N} p(v_i \mid S_1, S_2)}{\sum_{S_1,S_2} \prod_{i=1}^{N} p(v_i \mid S_1, S_2)}$$

Here again, The 2000-dimensional vector output $p(s_1|\mathbf{v})$ can be returned using the code in the Appendix. The visualisation of this posterior is plotted in Figure 2.

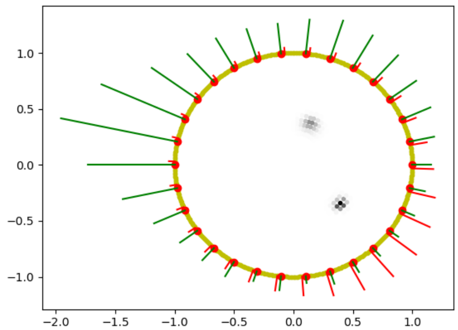


Figure 2: Plot of the maximum posterior $p(S_1 \mid \mathbf{v})$ for the scenario of two earthquakes with mean sensor data (darker corresponds to a higher probability). This is based on a spiral coordinate system with $\sigma$ = 0.2. The red points are the locations of the surface sensors. The observed (noisy) measurements at each sensor are represented by the green lines, and the true (unknown) blast values are denoted by the red lines.

**2.**

In this question, we proceed in a similar way as in question 2.1.2.

We are asked to report the value of $(log(p(v|H_2)) - log(p(v|H_1)))$.

Here again, for the set $H_2$, $v_i \sim N\left(\frac{0.5}{d_i^2(1)+0.1} + \frac{0.5}{d_i^2(2)+0.1}, \sigma^2\right)$ whilst, for the set $H_1$, $v_i \sim N\left(\frac{1}{d_i^2(1)+0.1}, \sigma^2\right)$,

as we are using the mean data set (EarthquakeExerciseMeanData.txt.)

Thus,

$$p(\mathbf{v} \mid H_2) = \sum_{S_1,S_2} p(\mathbf{v}, S_1, S_2 \mid H_2)$$

Assuming independence between S1 and S2:

$$\therefore p(\mathbf{v} \mid H_2) = \sum_{S_1,S_2} p(\mathbf{v} \mid S_1, S_2, H_2)p(S_1)p(S_2)$$

Likewise,

$$p(\mathbf{v} \mid H_1) = \sum_{S_1} p(\mathbf{v}, S_1 \mid H_1)$$

$$\therefore p(\mathbf{v} \mid H_1) = \sum_{S_1} p(\mathbf{v}, S_1 \mid H_1)p(S_1)$$

Notice that : $p(S_1) = p(S_2) = \frac{1}{2000}$

We use the model used in subquestion 2.2.1 to compute $p(\mathbf{v} \mid H_2)$ and create a new model based on the same spiral polar coordinate idea but with only one earthquake location S1 to compute $p(\mathbf{v} \mid H_1)$. The code to compute the posterior is in the Appendix and outputs **15.965927727784468**.

**3.**

The difference for this case is that sensors measure the mean values of incoming explosions. This leads to more homogeneity in the 30 sensor values and therefore decreases the relationship between an earthquake location and a sensor located close to the explosion. This results in a loss of information in the location of individual earthquakes by having more similar posterior probabilities across the locations and outputting a broader probable region as we can see by comparing Figure 2 with Figure 1.

By having mean data, you also lose magnitude relating to the number of explosions. The heterogeneity of the sensor values combined with their magnitudes provided information regarding the number of explosions. Indeed, having clear small regions of high probabilities makes it easier to estimate the number of locations. We realized in subquestion 1.3 that the estimation of the location directly relates to the estimation of the number of explosions. For section 1, the model was so sure it was not one explosion that the probability of it was approximated to zero therefore returning a log ratio of infinity. This clearly contrasts with the output of 15.97 for mean data which shows a substantial increase in difficulty to estimate the number of explosions.