ELEN3007A Group <u>27</u> - Assignment 2024:
Application of Bayes' Theorem for Locating a Robot's Position in an Enclosed Area

Kgadile E Masemola (876729)

September 13, 2024

QUESTION 1:

Why does θ_k not lie between $-\pi$ and π for which $p(\theta_k|\alpha, \beta, B)$ would then be $\frac{1}{2\pi}$?

Solution 1:

The given setup of the problem says that the photodetectors are placed on the x-axis above which the robot is located. Therefore, the signal comes from one(top) side of the axis(horizontal in this case). This thus limits the range of the detectors to only capture collimated flashes within the azimuth $-\frac{\pi}{2}$ to $\frac{\pi}{2}$. .

QUESTION 2:

Prove that

$$p(x_k|\alpha, \beta, B) = \frac{\beta}{\pi(\beta^2 + (x_k - \alpha)^2)} \tag{1}$$

Solution 2:

Using the relation $x_k = \alpha + \beta \tan \theta_k$ (given in the brief [1]) and the prior PDF $p(\theta_k|\alpha, \beta, B) = f_{\Theta|\Lambda,\beta,B}(\cdot) = \frac{1}{\pi}$ and the transformation of random variables (notations for random variables defined **Solution 8**). The change of variables $x = g(y) \Rightarrow \theta_k = g(x_k)$. The function $f(\theta)$ becomes $\tilde{f} = f(g(x))$. The given PDF $f_\Theta(\theta)$ corresponds to $f_X(x)$ with respect to new variable x. Therefore change of variables for PDF's equation adapted from [2] is:

$$p_x(x) = p_\theta(\theta)\left|\frac{d\theta}{dx}\right| = p_\theta(g(x))|g^{\cdot}(x)| \tag{2}$$

using the correct notation and given x_k relation:

$$f_X(x_k) = f_\Theta(\theta_k) \times \left|\frac{d\theta_k}{dx_k}\right| = f_\Theta(\theta_k) \times \left|\left(\frac{dx_k}{d\theta_k}\right)^{-1}\right| \tag{3}$$

$$= \frac{1}{\pi} \times \left|\left(\frac{d}{d\theta_k}(\beta\tan\theta_k + \alpha)\right)^{-1}\right| = \frac{1}{\pi} \times \left|\frac{1}{\beta\sec^2\theta_k}\right|$$

$$but \ \sec^2(\theta_k) = 1 + \tan^2\theta_k \ and \ \tan\theta_k = \frac{x_k - \alpha}{\beta} \ (relation) \tag{4}$$

$$so \ = 1 + \left(\frac{x_k - \alpha}{\beta}\right)^2 = \frac{\beta^2 + (x_k - \alpha)^2}{\beta^2}$$

$$\therefore \ f_X(x_k) = \frac{1}{\pi} \times \frac{\beta^2}{\beta(\beta^2 + (x_k - \alpha)^2}$$ (5)

$$and, \ so \ f_X(x_k) \leftrightharpoons f_{X|\Lambda,\mathfrak{B},B}(\cdot \,|\alpha, \beta) \ and \ f_\Theta(\theta_k) \leftrightharpoons f_{\Theta|\Lambda,\beta,B}(\cdot)$$

$$thus \ f_{X|\Lambda,\mathfrak{B},B}(\cdot \,|\alpha, \beta) = \frac{\beta}{\pi(\beta^2 + (x_k - \alpha)^2)} \ \blacksquare$$ (6)

QUESTION 3:

Plot $p(x_k|\alpha, \beta, B)$ and then relate its width at half maximum to the parameters of the PDF.

Solution 3:

The plot of the PDF $p(x_k|\alpha, \beta, B) \Rightarrow f_{X|\Lambda,\mathfrak{B},B}(x_k|\Lambda = \alpha, \mathfrak{B} = \beta)$ is shown in Figure 1 below. The *parameters* are chosen to be $\alpha = 4; \beta = 2$ for $\{x_k\}_{k=-10}^{10}$. It shows that the maximum value is when $x_k = 4 = \alpha$. At half maximum of *these* paramaters, the width at half maximum of the distribution is $4 = 2\beta$, half maximum height is $y = 0.1591536$.
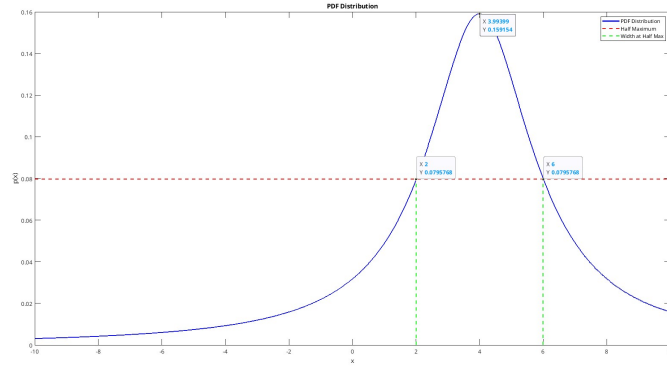

Figure 1: Q3 Plot of the PDF $p(x_k|\alpha, \beta, B)$

QUESTION 4:

Derive the expression for $p(\alpha|x_k, \beta, B)$.

Solution 4:

The conditional distribution of $\Lambda(\cdot)$ given $X(\cdot)$, $\mathfrak{B}(\cdot)$, Bayes' Theorem requires equation (7) *(notation in **Solution 8**)*, from [3]:

$$f_{\Lambda|X,\mathfrak{B},B}(\cdot \,|x_k, \beta) = \frac{f_{X|\Lambda,\mathfrak{B},B}(\cdot \,|x_k, \beta)}{f_X(x_k)} = \frac{f_{X|\Lambda,\mathfrak{B},B}(\cdot \,|x_k, \beta)f_\Lambda(\cdot)}{f_X(x_k)}$$ (7)

But, we can similarly use the same method as proof in **Solution 2** by setting α as the subject of the formula in the relation $\alpha = -(\beta \tan \theta_k - x_k)$. We get:

2

$$f_{\Lambda|X,\mathfrak{B},B}(\,\cdot\,|x_k,\beta) = f_\Theta(\theta_k) \times \left|\frac{d\theta_k}{d\alpha}\right| = f_\Theta(\theta_k) \times \left|\left(\frac{d\alpha}{d\theta_k}\right)^{-1}\right| \tag{8}$$

$$but, \;\; \frac{d\alpha}{d\theta} = -\beta \sec^2\theta_k \;\; and \;\; \left|\left(\frac{d\alpha}{d\theta_k}\right)^{-1}\right| = \frac{1}{\beta \sec^2\theta_k} \tag{9}$$

as above, substituting (9) in (8) and $f_\Theta(\theta_k) = \frac{1}{\pi}$, and using **2**, we get still get:

$$f_{\Lambda|X,\mathfrak{B},B}(\,\cdot\,|x_k,\beta) = \frac{1}{\pi}\frac{\beta}{(\beta^2 + (x_k - \alpha)^2)} \;\; \blacksquare \tag{10}$$

QUESTION 5:

Finally derive an expression for $p(\alpha|\{x_k\}_{k=1}^N, \beta, B)$. (State all assumptions.)

Solution 5:

Assumptions:

- The prior PDF is given as $f_{\Theta|\Lambda,\mathfrak{B},B}(\cdot) = \frac{1}{\pi}$. And using the transformation of random variables in **2** becomes equation (6).

- The measurements $\{x_k\}_{k=1}^N$ are input to the random variable $X_k(\cdot) = \{x_k\}_{k=1}^N$

- $\{x_k\}_{k=1}^N$ are independent, and continous.. thus $X_k(\cdot)$ is a continuous random variable.

- The normalisation/marginal probability of observed data/ proportionality factor $f_X(x_k) \approx$ *weird integral* $\simeq$ *constant* in eqaution (7) is discarded $\because$ it does not depend on α. It just a scalling factor which doesn't affect the peak point α.

- So we can use the Posterior density $\propto$ Likelyhood $\times$ Prior Density [3], where the distributions are infered in (10):

$$f_{\Lambda|X,\mathfrak{B},B}(\,\cdot\,|x_k,\beta) \propto f_{X|\Lambda,\mathfrak{B},B}(\,\cdot\,|\alpha,\beta)f_\Lambda(\alpha) \tag{11}$$

Then the joint distribution of $p(\alpha|\{x_k\}_{k=1}^N, \mathfrak{B}, B)$:

$$f_{\Lambda|X_k,\mathfrak{B},B}(\alpha|x_1,...,x_N,\beta) = \prod_{k=1}^N \frac{\beta}{\pi(\beta^2 + (x_k - \alpha)^2)} \times \aleph \;\; \blacksquare \tag{12}$$

$\aleph$ is the proportionality factor which may be ignored.

QUESTION 6:

Explain how one obtains the x-position of the robot from the result in (5.).

Solution 6:

The result in 5. can be used to obtain the position x-position of the robot by taking the product of the distributions of each x_k measurement. This is due to the fact that the measurements are independent they are independent.

The x-position can also be obtained by locating the peak of the distribution (plot) and mapping it to the random variable variable axis. Additionally another method is to use Maximum-A-Posterior (MAP). We can also estimate the posterior mean by applying maximum likelyhood solution μ_{ML} to the mean of the data distribution and the mean of the prior distribution μ_0. This way the estimate will be a compromise between μ_0 and μ_{ML}. The more measurements we use, the more the estimate approximate to μ_{ML}.

QUESTION 7:

Implement the Bayesian position inference scheme in Matlab. Assume the robot is located inside a confined square region of size 10m × 10m. Demonstrate your Bayes estimator by inferring the robot's x-position from the data/measurements $\{x_k\}_{k=1}^{200}$ in *BayesData.mat*, with the robot's y-position 4.5 m. Plot the Bayes posterior distribution for the first $N = 1, 2, 5$ and 30 measurements. For this data, what is your best x-position estimate?

Solution 7:

The Bayesian position inference scheme MATLAB code is attached in Appendix A. The plots are attached in Appendix B as Figure 2 & 3. The scheme tests the first $N = 1, 2, 5, and$ 30. Case 1 is for the first points inside that are detected inside the 10×10 box are considered. Case 2 does not have restrictions and this uses the first values.

It is noticed that when points outside the box, the estimate $\alpha \approx 5.85$ which is within the box at, close the centre. However, when all points are considered, the mean estimates $\alpha \approx 10.16$ for $N = 30$. But strangely this is outside the box. In comparing the distribution for the different N-measurements, the distribution seems to move as more measurements are known.

Maximum likelyhood estimate (MLE) solution was applied to the data $\{x_k\}_{k=1}^{200}$, attached in Appendix C, and the estimation of $\alpha \approx 7.2287$. Comparing this value to posterior-PDF values inside the box (Figure 2) for N = 2 & 30, the approximate value of α moves further away from the MLE value.

After conducting this assignment, we conclude that: 1) The posterior distribution provides a way to quantify the uncertainty in the robot's position estimate. Instead of providing a single "best guess," Bayesian inference offers a probabilistic distribution over all possible positions, allowing for a richer representation of uncertainty.; 2) The number of measurements significantly affects the shape of the posterior distribution. The more data we have, the better your estimates; 3) In contrast to maximum likelihood estimation (MLE), which finds the single parameter value that maximizes the likelihood, Bayesian inference provides a full posterior distribution, reflecting a range of possible outcomes.

After this analyis, the best estimate of α is ≈ 5.85568 as obtained from Posterior Infererence for points inside the box and $N = 30$. See Figure 2, Appendix B. ∎

QUESTION 8:

The notation presented above, deliberately does *not* follow the conventions your Probs lecturer introduced. Throughout your assignment, strictly use the notation prescribed for use in Probs.

Solution 8:

Yes with the following Random Variable Notations and Equations:

$$\Lambda(\cdot) = \alpha; \; \mathfrak{B}(\cdot) = \beta; \; X(\cdot) = x_k; \; \Theta(\cdot) = \theta_k \tag{13}$$

$$p(\theta_k|\alpha,\beta,B) = f_{\Theta|\Lambda,\beta,B}(\,\cdot\,|\alpha,\beta) = f_{\Theta|\Lambda,\beta,B}(\cdot) \tag{14}$$

$$p(\alpha|x_k,\beta,B) = f_{\Lambda|X,\mathfrak{B},B}(\,\cdot\,|x_k,\beta) = f_{\Lambda|X,\mathfrak{B},B}(\cdot) \tag{15}$$

$$p(x_k|\alpha,\beta,B) = f_{X|\Lambda,\mathfrak{B},B}(\,\cdot\,|\alpha,\beta) \tag{16}$$

QUESTION 9:

Professional report with clear and effective data representation. The report must not have a title page and is not allowed to exceed 5 pages in length.

Solution 9:

The report is presented in homework assignment form(statement and solution). Additionally, the following reference materials have been consulted for the assignment: [4], [2]

(Bonus) QUESTION 10:

Infer both the x-position as well as the y-position of the robot. Estimate both α and β for the PDF in Eq. (1).

Solution 10:

References

[1] Prof. Anton van Wyk. Application of bayes' theorem for locating a robot's position in an enclosed area. *School of Electrical and Information Engineering - Wits University*. ELEN3007A - Assignment Brief, 2024.

[2] Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*, volume 4. Springer, 2006.

[3] Bradley Efron. Lecture notes for stats 200: Introduction to statistical inference. https://web.stanford.edu/class/archive/stats/stats200/stats200.1172/Lecture20.pdf, 2017. Accessed: 2024-09-12.

[4] Alberto Leon-Garcia. Probability, statistics, and random processes for electrical engineering. 2017.

Appendix A. - Matlab Inference Scheme Code

```matlab
1   %%%%%%%%Bayesian Position Inference Scheme%%%%%%%%%%%%
2   %numbat24.04%
3   %
4   % Main function to execute the Bayesian inference and plot results
5   function BayesianPositionInferenceScheme()
6       load('BayesData.mat');
7       % Given beta value and alpha range
8       beta = 4.5;
9       alpha_values = linspace(-50, 50, 1000);
10
11      % Number of measurements to evaluate the posterior
12      N_values = [1, 2, 5, 30];
13
14      % Case 1: Points Inside the Box (0 <= x <= 10)
15      x_inside_box = x(x >= 0 & x <= 10);
16      plot_posterior(x_inside_box, N_values, alpha_values, beta, ...
17          'Posterior for Points Inside the Box (0 <= x <= 10)');
18
19      % Case 2: All Points (No Restrictions)
20      x_all_points = x;
21      plot_posterior(x_all_points, N_values, alpha_values, beta,
22      'Posterior for All Points');
23  end
24
25  % Function to calculate the posterior for a given dataset and N points
26  function posterior_values = calculate_posterior(data, N, alpha_values, beta)
27      posterior_values = zeros(size(alpha_values)); % Initialize posterior array
28      for i = 1:length(alpha_values)
29          alpha = alpha_values(i); % Current alpha value
30          likelihood = 1; % Initialize likelihood as 1
31          for k = 1:min(N, length(data))
32              x_k_val = data(k);
33              % Calculate the likelihood for each x_k_val
34              likelihood = likelihood*(beta / (pi * (beta^2 + (x_k_val - alpha)^2)));
35          end
36          posterior_values(i) = likelihood; % Store posterior value
37      end
38      % Normalize the posterior to sum to 1
39      posterior_values = posterior_values / sum(posterior_values);
40  end
41
42  % Function to plot the posterior for different N values
43  function plot_posterior(data, N_values, alpha_values, beta, title_text)
44      figure; % Create new figure
45      hold on;
46      for N = N_values
47          % Calculate posterior for the current N
48          posterior_values = calculate_posterior(data, N, alpha_values, beta);
49          % Plot the posterior distribution for this N
50          plot(alpha_values, posterior_values, 'DisplayName', ['N = ', num2str(N)]);
51      end
52      xlabel('Alpha');
53      ylabel('Posterior Probability');
54      title(title_text);
55      legend show;
56      grid on;
57      hold off;
58  end
```
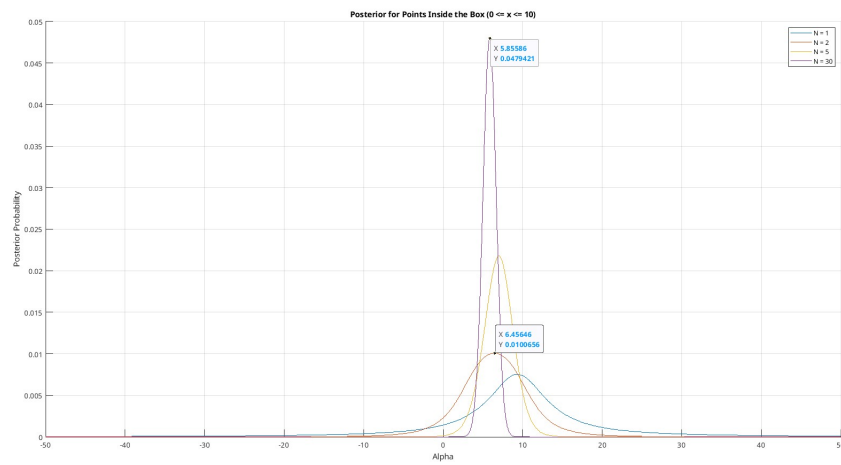
Appendix B. - Matlab Scheme Plots

Figure 2: Bayes' Inference Scheme Plot of the PDF for points inside the box
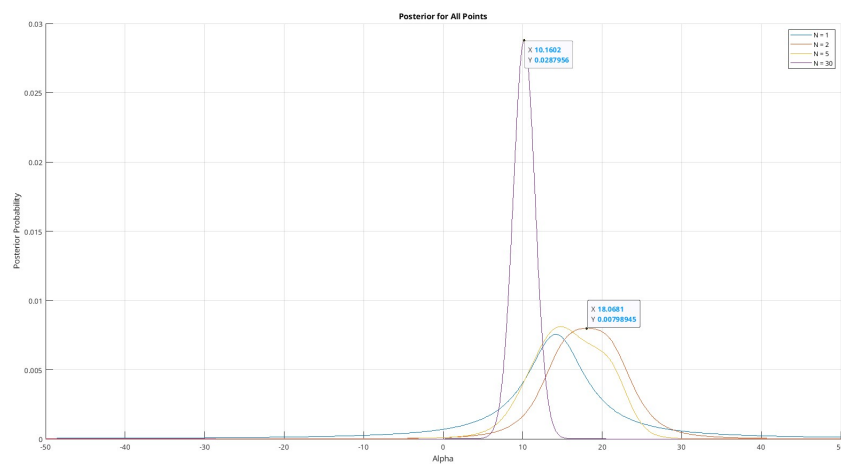

Figure 3: Bayes' Inference Scheme Plot of the PDF for all points

Appendix C. - Maximul Likelyhood Estimation Scheme

```matlab
1   %%%%%%%Estimating alpha using Maximum Likelyhood Solution%%%%%%%%%%%
2   %numbat24.04%
3   %
4   % Load the data from BayesData.mat
5   load('BayesData.mat');
6
7   % Given y-position (beta)
8   beta = 4.5;
9
10  % Define the log-likelihood function so that product becomes a sum
11  log_likelihood = @(alpha) sum(log(beta ./ (pi * (beta^2 + (x - alpha).^2))));
12
13  % Find the alpha that maximizes the log-likelihood,
14  % use fminsearch to minimize the negative log-likelihood, (equivalent to
15  % maximizing the likelihood)
16  alpha_mle = fminsearch(@(alpha) -log_likelihood(alpha), mean(x));
17
18  % Display the MLE result
19  disp(['Maximum Likelihood Estimate for alpha: ', num2str(alpha_mle)]);
20  %% alpha = 7.2287 %%
```