AER1513 Assignment 1

Yeqi Sang
1002748681
abel.sang@mail.utoronto.ca

October 3, 2024

Question 1

The assumption of zero-mean Gaussian noise is reasonable. According to Figure 1.6, for range error:

- Mean: $-0.000000\,\mathrm{m}$, StdDev: $0.019155\,\mathrm{m}$.
- For speed error: Mean: $-0.000451\,\mathrm{m/s}$, StdDev: $0.047554\,\mathrm{m/s}$.

The value of the speed error mean is small enough to be considered zero-mean. Since w_k is the noise of speed measurement and n_k is the noise of observation (laser range measurement):

$$\sigma_q^2 = 0.047554^2 = 0.00226\,\mathrm{m}^2/\mathrm{s}^2$$

for the speed measurements and

$$\sigma_r^2 = 0.019155^2 = 0.000367\,\mathrm{m}^2$$

for the range measurements.

Question 2

In MAP estimation, we want to maximize the posterior:

$$x^* = \arg\max P(x|y, u) = \arg\max P(y|x, u)P(x|u)$$

Since $P(y|u)$ is constant and we can take logarithm of the equation to:

$$x^* = \arg\max[\log P(y|x) + \log P(x|u)]$$

The motion model is:

$$x_k = x_{k-1} + Tu_k + w_k$$

where w_k is Gaussian. Thus, the prior probability is:

$$P(x|u) \propto \exp\left(-\sum_{k=1}^{K} \frac{(x_k - (x_{k-1} + Tu_k))^2}{2\sigma_q^2}\right)$$

The range measurements are modeled by:

$$y_k = x_k + n_k$$

where n_k is Gaussian. Thus, the likelihood is:

$$P(y|x) \propto \exp\left(-\sum_{k=1}^{K} \frac{(y_k - x_k)^2}{2\sigma_r^2}\right)$$

Then to simplify the equation that we want to maximize, we can get:

$$J(x_{1:K}|u_{1:K}, y_{1:K}) = \sum_{k=1}^{K} \frac{(x_k - (x_{k-1} + Tu_k))^2}{2\sigma_q^2} + \sum_{k=1}^{K} \frac{(y_k - x_k)^2}{2\sigma_r^2}$$

Question 3

To get the optimal position estimates, we need to make $J(x_{1:K}|u_{1:K}, y_{1:K})$ as a quadratic equation:

$$J(x|u,y) = \frac{1}{2}(x_k - A_{k-1}x_{k-1} - Tu_k)^T Q_k^{-1}(x_k - A_{k-1}x_{k-1} - Tu_k) + \frac{1}{2}(y_k - x_k)^T R_k^{-1}(y_k - x_k)$$

Thus:

$$J(x|u,y) = \frac{1}{2}(z - Hx)^T W^{-1}(z - Hx)$$

where

$$z = \begin{pmatrix} Tu_1 \\ \vdots \\ Tu_2 \\ y_1 \\ \vdots \\ y_k \end{pmatrix}$$

and

$$H = \begin{pmatrix} A^{-1} \\ C \end{pmatrix}$$

For the components in H:

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}$$

Thus

$$A^{-1} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ -1 & 1 & \dots & 0 & 0 \\ 0 & -1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 1 & \vdots \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

and

$$C = \text{diag}(C_1, C_2, \dots, C_k) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

The matrix Q is a diagonal matrix, where each diagonal element is σ_q^2, and the size of the matrix is $k \times k$:

$$Q = \text{diag}(\sigma_q^2, \sigma_q^2, \dots, \sigma_q^2) = \begin{pmatrix} \sigma_q^2 & 0 & \dots & 0 \\ 0 & \sigma_q^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_q^2 \end{pmatrix}$$

Similarly, the matrix R is a diagonal matrix, where each diagonal element is σ_r^2, and the size of the matrix is $k \times k$:

$$R = \text{diag}(\sigma_r^2, \sigma_r^2, \dots, \sigma_r^2) = \begin{pmatrix} \sigma_r^2 & 0 & \dots & 0 \\ 0 & \sigma_r^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r^2 \end{pmatrix}$$

Finally, the combined matrix R that includes both Q and R in a block diagonal form can be represented as:

$$W = \text{diag}(Q_1, Q_2, \dots, Q_K, R_1, R_2, \dots, R_k) = \begin{pmatrix} Q_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & Q_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Q_K & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & R_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & R_k \end{pmatrix}$$

Taking the partial derivative of $J(x|u, y)$ with respect to x and setting it to zero gives:

$$(H^T W^{-1} H)\hat{x} = H^T W^{-1} z$$

Thus, we obtain the optimal position estimates $\hat{x}$:

$$\hat{x} = (H^T W^{-1} H)^{-1} H^T W^{-1} z$$

Question 4

Based on the property of A and A^{-1}, the sparsity comes from the fact that the system model from Question 3 obeys the Markov property. When we plug A^{-1} back into the linear system from Question 3, using the data we have, the left-hand side $(H^T W^{-1} H)$ looks like:

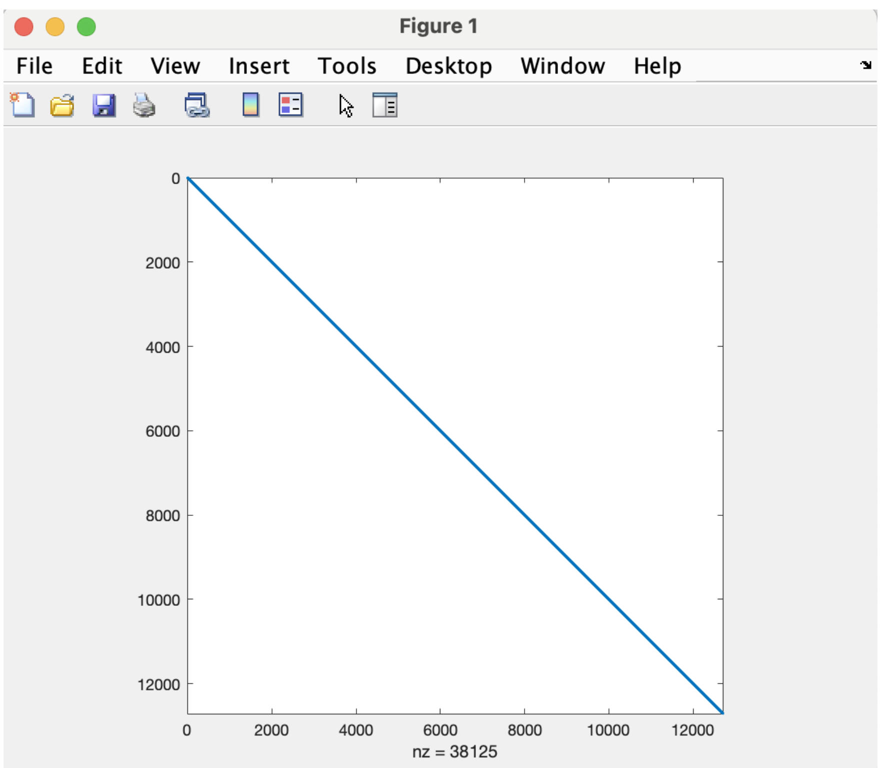


Figure 1: Sparsity Pattern based on the dataset.

We can then apply sparse Cholesky factorization, which gives:

$$H^T W^{-1} H = LL^T$$

Finally, we can solve for d, where:

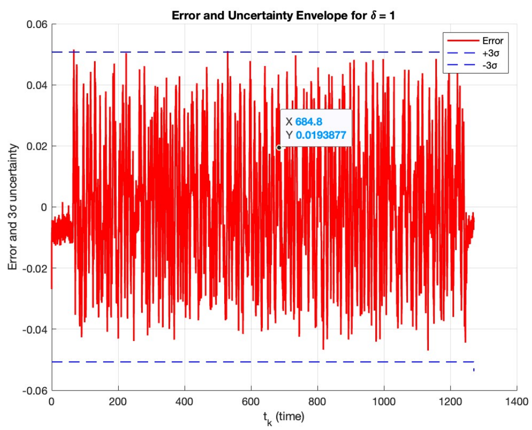
$$Ld = H^T W^{-1} z$$

and

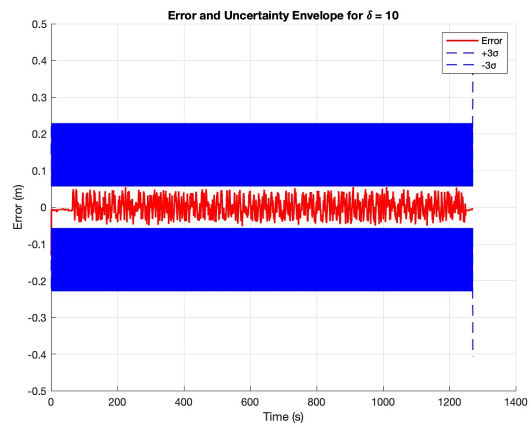
$$d = L^T x$$

Then We can solve for $\hat{x}$ in $O(N^3 K)$ instead of $O(K^3)$, where N is the time complexity of the Cholesky factorization, which is much smaller than K in our case.

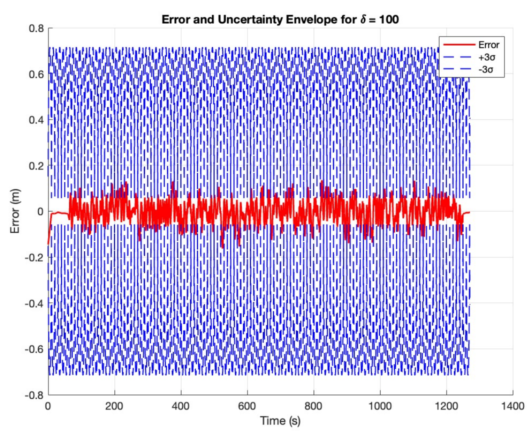
Question 5



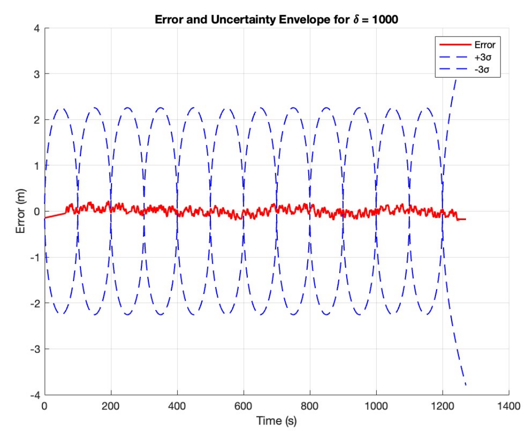
(a) Error for $\delta = 1$



(b) Error for $\delta = 10$

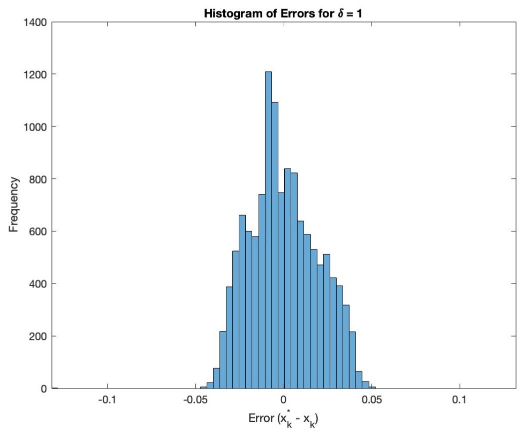


(c) Error for $\delta = 100$

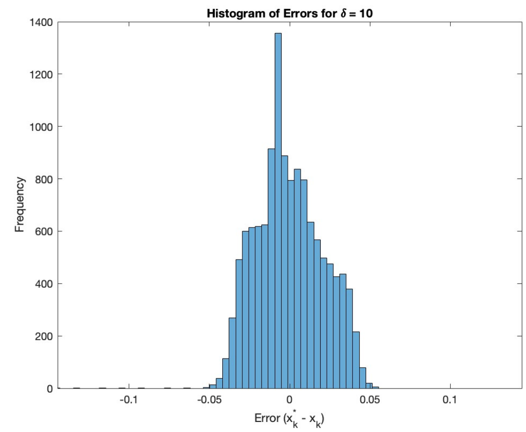


(d) Error for $\delta = 1000$

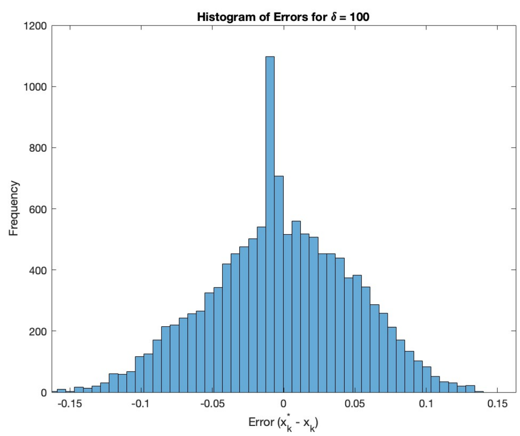
Figure 2: Error and Uncertainty Envelope for different δ values.



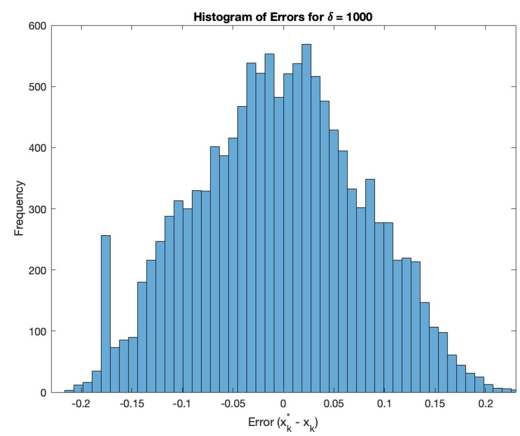
(a) Errors for $\delta = 1$



(b) Errors for $\delta = 10$



(c) Errors for $\delta = 100$



(d) Errors for $\delta = 1000$

Figure 3: Histograms of Errors for different δ values.

As δ increases, the uncertainty envelope are wider, showing that less frequent measurements lead to greater uncertainty in the estimator's predictions. The error remains close to zero for all δ, but with increasing variability as the estimator relies more on predictions and fewer corrections. In the histograms, errors are tightly concentrated around zero for $\delta = 1$, but as δ increases, the spread of errors grows, reflecting less precision. Overall, larger δ values lead to higher uncertainty and more error variability, though the estimator remains centered around the true value.

A Appendix

Listing 1: Matlab code

```
load('dataset1.mat');
who

K = size(t, 1);

estimates_data = zeros(K, 4);

deltas = [1, 10, 100, 1000];

T = 0.1;
d = 3;
for d = 1:length(deltas)
    delta = deltas(d);

    A = tril(ones(K));

    inv_A = eye(K);

    for i = 2:K
        inv_A(i, i-1) = -1;
    end

    C = speye(K);
    C = C(1:delta:end, :);

    H = [inv_A; C];
    H = sparse(H);

    y = l - r;
    y = y(1:delta:end, :);

    observation_size = size(y, 1);

    z = [T*v; y];

    Q = v_var * eye(K);
    R = r_var * eye(observation_size);

    inv_Q = 1/v_var * eye(K);
    inv_R = 1/r_var * eye(observation_size);

    W = [Q, zeros(K, observation_size); zeros(observation_size,K), R];
    W = sparse(W);

    inv_W = [inv_Q, zeros(K, observation_size); zeros(observation_size,K), inv_R];
    inv_W = sparse(inv_W);

    left_side = H' * inv_W * H;
    left_side = sparse(left_side);
    L = chol(left_side, 'lower');

    d_d =  L \ (H' * inv_W * z);

    x_estimate = L' \ d_d;

    estimates_data(:, d) = x_estimate;


    P_inv = L' \ (L \ eye(K));
    variance_x_estimates(:, d) = diag(P_inv);
end
```

```
for d = 1:length(deltas)

    delta = deltas(d);

    error = estimates_data(2:end, d) - x_true(2:end);

    sigma_xk = sqrt(variance_x_estimates(2:end, d));
    upper_bound = 3 * sigma_xk;
    lower_bound = -3 * sigma_xk;

    t_plot = t(2:end);

    figure;
    hold on;


    plot(t_plot, error, 'r-', 'LineWidth', 1.5);
    % Plot uncertainty envelope as dotted blue lines
    plot(t_plot, upper_bound, 'b--', 'LineWidth', 1);
    plot(t_plot, lower_bound, 'b--', 'LineWidth', 1);

    grid on;

    xlabel('Time (s)');
    ylabel('Error (m)');
    title(['Error and Uncertainty Envelope for \delta = ', num2str(delta)]);
    legend('Error', '+3  ', '-3  ');

    hold off;
end


for d = 1:length(deltas)
    delta = deltas(d);
    error = estimates_data(:, d) - x_true;

    figure;
    histogram(error, 50);

    % Adjust x-axis to center at 0
    max_error = max(abs(error));
    xlim([-max_error, max_error]);

    xlabel('Error (x^*_k - x_k)');
    ylabel('Frequency');
    title(['Histogram of Errors for \delta = ', num2str(delta)]);

    hold off;
end
```