Problem 1

A system satisfies the partial differential equation

$$\ddot{z} + 10\dot{z} + 100z = f(t), \quad z(0) = \dot{z}(0) = 0$$

where $z(t)$ is the response variable (e.g. position) and $f(t)$ is a white noise forcing function with a PSD of 10 'units'.

part (a)

Let z and $\dot{z}$ be the state variables for this system. Write out the dynamics for this CT system in stochastic LTI form.

If $x = [z, \ \dot{z}]^T$, then the system in LTI form is expressed as

$$\dot{x} = \begin{bmatrix} \dot{z} \\ \ddot{z} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -100 & -10 \end{bmatrix} \begin{bmatrix} z \\ \dot{z} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} f(t)$$

part (b)

Suppose this process is sampled at a uniform rate beginning at time $t = 0$, with $\Delta T = 0.2$ seconds. Convert the CT stochastic LTI model from part (a) into DT form, and report the corresponding F and Q matrices.

$$F = \begin{bmatrix} 0.151 & 0.042 \\ -4.193 & -0.269 \end{bmatrix} \qquad Q = \begin{bmatrix} 0.004 & 0.009 \\ 0.009 & 0.376 \end{bmatrix}$$

part (c)

Suppose the units of z are in meters and $\dot{z}$ are in m/s. Provide the corresponding units of the elements of Q, and explain what Q says physically about the elements of the DT process noise vector $w(k)$ and its effect on the DT state $x(k)$.

If x_1 is in m and x_2 is in m/s, then the units of $Q_{1,1}$ are in m^2 and $Q_{2,2}$ are in (m/s)^2. Because Q is the covariance of the multivariate white gaussian noise distribution, the diagonal elements of Q represent the variances of the additive noise to the state vector and the off-diagonal elements represent their covariances.

part (d)

Suppose a sensor is attached to this system, with the CT model $y(t) = z + \dot{z}\Delta T + \tilde{v}(t)$, where $\tilde{v}(t)$ is a white noise input with PSD of 3 units. Report the H and R matrices for the corresponding DT model; also report the units of R if z has units of meters.

$$H = [1 \ \Delta T]^T$$

Because $y(k)$ is a scalar, the R 'matrix' is simply the 1D variance of the additive white gaussian noise. This means $R = 3$.

Problem 2

Consider a battery with a completely unknown voltage ($P_0 = \infty$). Two independent measurements of the voltage are taken to estimate the voltage, the first with a variance of 1 and the second with a variance of 4.

part (a)

Write the weighted least squares voltage estimate in terms of the two measurements y_1 and y_2.

The measurement model for this system is $y(k) = Hx + \mathcal{N}(0, R)$ where $y = [y_1, \ y_2]^T$, $H = [1, \ 1]^T$, and

$$R = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

So the weighted least squares estimate of this system's state is

$$\hat{x}_{LS} = (H^T R^{-1} H)^{-1} H^T R^{-1} y$$

$$= \left\{ [1 \ 1] \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}^{-1} [1 \ 1] \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= 1.25^{-1} [1 \ 0.25] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= \frac{y_1 + 0.25 y_2}{1.25}$$

part (b)

If weighted least squares is used to estimate the voltage, what is the variance of the voltage estimate after the first measurement? After the second measurement?

The covariance of the estimate is calculated as $P_e = (H^T R^{-1} H)^{-1}$. The covariance after the first measurement is simply $(1 \times 1^{-1} 1)^{-1} = 1$. The covariance after the second measurement is

$$\left\{ [1 \ 1] \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}^{-1} = 0.8$$

part (c)

If the voltage is estimated as $(y_1 + y_2)/2$, what is the variance of the voltage estimate?

In this case $\hat{x}$ is estimated as $(H^T H)^{-1} H^T y$. This means the estimation error $e = -(H^T H)^{-1} H^T v$. Because of this the estimate covariance becomes

$$
\begin{aligned}
P_e &= E(ee^T) \\
&= E((H^T H)^{-1} H^T v \{\ldots\}^T) \\
&= (H^T H)^{-1} H^T E(vv^T)(H^T)^T (H^T H)^{-T} \\
&= (H^T H)^{-1} H^T R H (H^T H)^{-T} \\
&= \frac{1}{2} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{1}{2} \\
&= 1.25
\end{aligned}
$$

Problem 3

The production of steel in the United States between 1946 and 1956 was 66.6, 84.9, 88.6, 78.0, 96.8, 105.2, 93.2, 111.6, 88.3, 117.0, and 115.2 million tons. Find the least squares fit to these data using the following functions. Include a plot of the original data along with the least squares curve, the rms error of the least squares fit, and a prediction for steel production in 1957.

part (a)

Linear Curve Fit

To find the linear fit coefficients we model each measurement as $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ where β_0 and β_1 are the coefficients and ϵ_i is the noise at timestep i. This model is expressed in matrix form as

$$
\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}
$$

or $Y = X\beta + \epsilon$. This equation is solved for β as $\hat{\beta} = (X^T X)^{-1} X^T Y$, giving us the linear fit pictured below in figure 1. The RMS error for the linear fit is 8.7823 and the predicted steel production for 1957 is 118.715 million tons.

part (b)

Quadratic Curve Fit

To find the polynomial fit coefficients we simply extend the linear model as $y_i = \beta_0 + \beta_1 x_i + \cdots + \beta_m^m + \epsilon_i$ where β_i are the coefficients of a polynomial of degree m and ϵ_i is the noise at timestep i. This model is expressed in matrix form as

$$
\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^m \\ 1 & x_1 & x_1^2 & \cdots & x_1^m \\ \vdots & \vdots & \ddots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \cdots \\ \beta_m \end{bmatrix} + \begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}
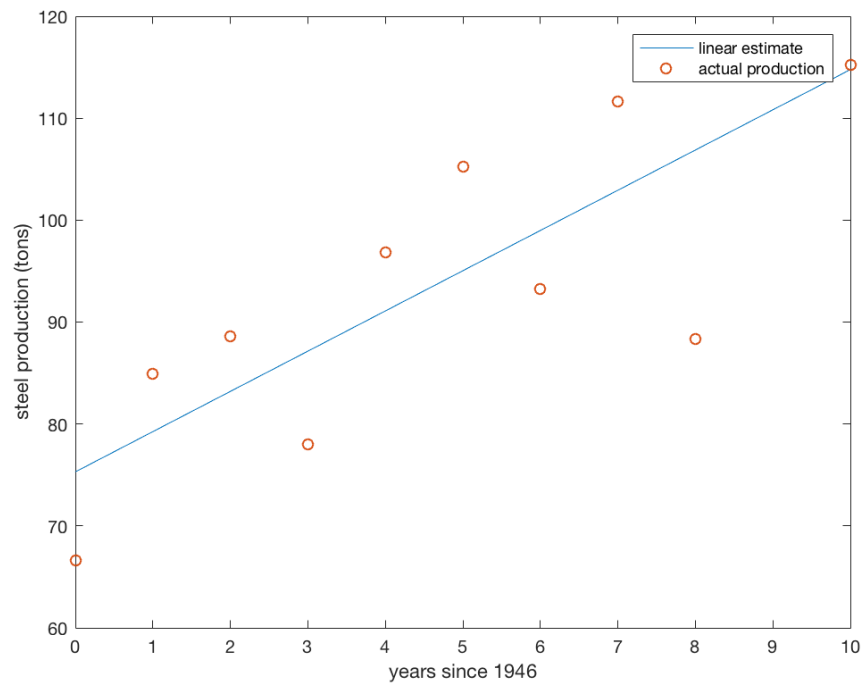$$

Figure 1: estimated vs actual state for linear fit

or $Y = X\beta + \epsilon$. We can still solve for β as $\hat{\beta} = (X^T X)^{-1} X^T Y$, giving us the least squares approximation for the polynomial coefficients. Note that parts (c) and (d) used this same method but with $m = 3$ and 4, respectively. The quadratic (degree 2 polynomial) fit is pictured below in figure 2. The RMS error for the linear fit is 8.6665 and the predicted steel production for 1957 is 114.530 million tons.

part (c)

Cubic Curve Fit

The RMS error for the cubic fit is 8.2889 and the predicted steel production for 1957 is 126.188 million tons. Figure 3 shows the estimated vs actual steel production for 1946 through 1956.

part (d)

Quartic Curve Fit

The RMS error for the quartic fit is 8.2816 and the predicted steel production for 1957 is 128.861 million tons. Figure 4 shows the estimated vs actual steel production for 1946 through 1956.
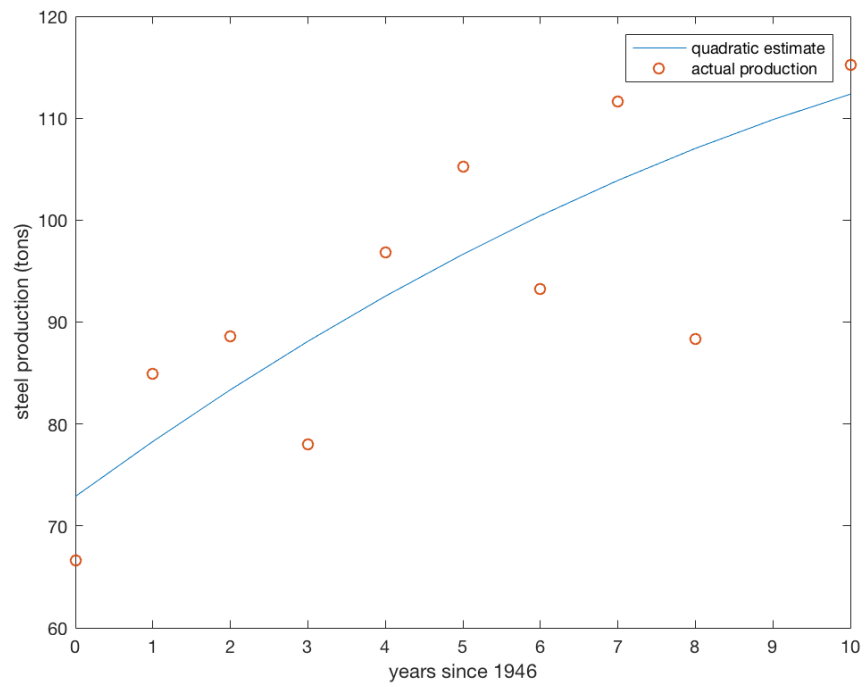
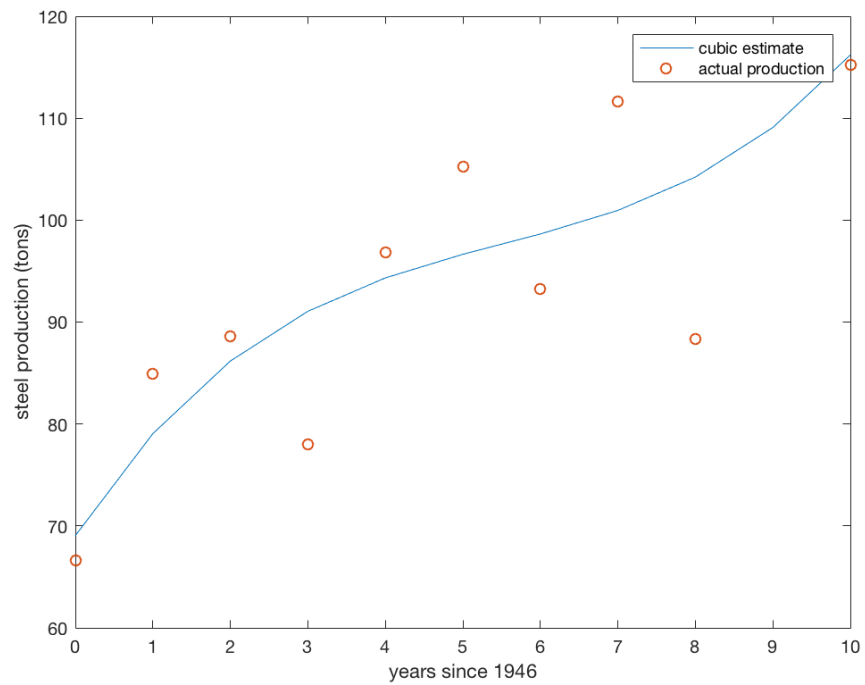Figure 2: estimated vs actual state for quadratic fit

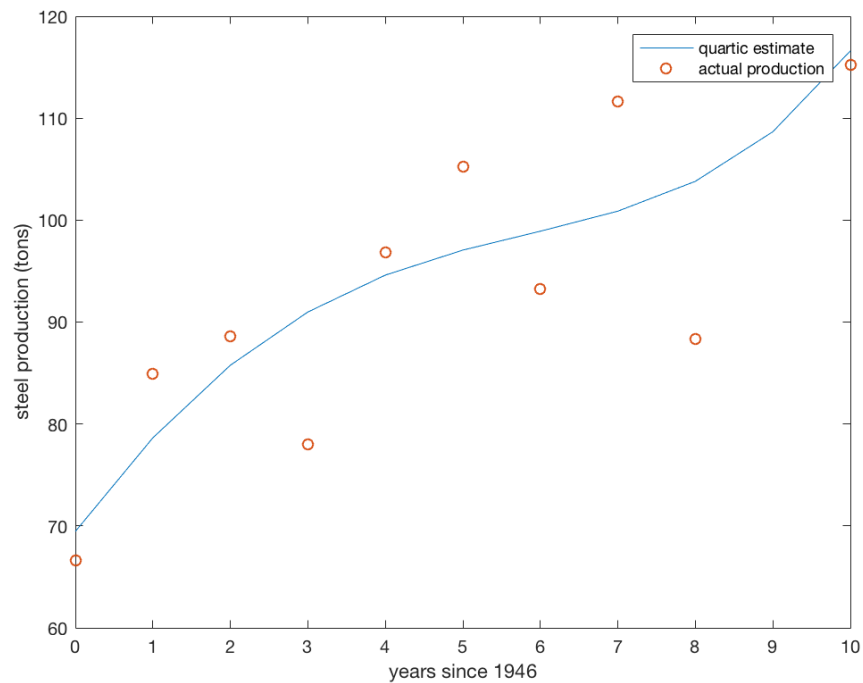Figure 3: estimated vs actual state for cubic fit

Figure 4: estimated vs actual state for cubic fit

Problem 4

Consider the 3D static robot GPS localization problem defined in Lecture 19. Suppose the AWGN sensor error covariance matrix is given by

$$R = \begin{bmatrix} 8 & 5.15 & 6.5 \\ 5.15 & 5 & -4.07 \\ 6.5 & -4.07 & 50 \end{bmatrix}$$

part (a)

Suppose the true robot position is known to be at $\xi_0 = \eta_0 = z_0 = 1\text{m}$. Simulate $T = 100$ measurements for the robot at this location and provide three separate 2D scatter plots of the resulting simulated noisy y data: one showing all $y_k(1)$ vs $y_k(2)$, another showing all $y_k(1)$ vs $y_k(3)$, and another showing all $y_k(2)$ vs $y_k(3)$. Use the same scales on both axes for all of the plots.

part (b)

Compute the sample covariance matrix for the simulated measurement vectors from part (a). How does the estimated covariance compare to R?
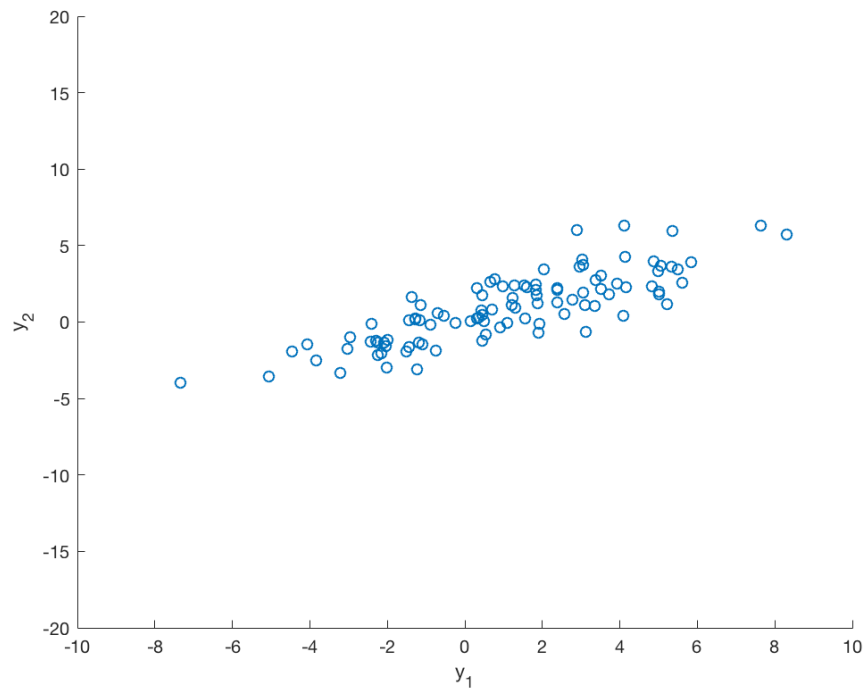
Figure 5: y_1 vs y_2
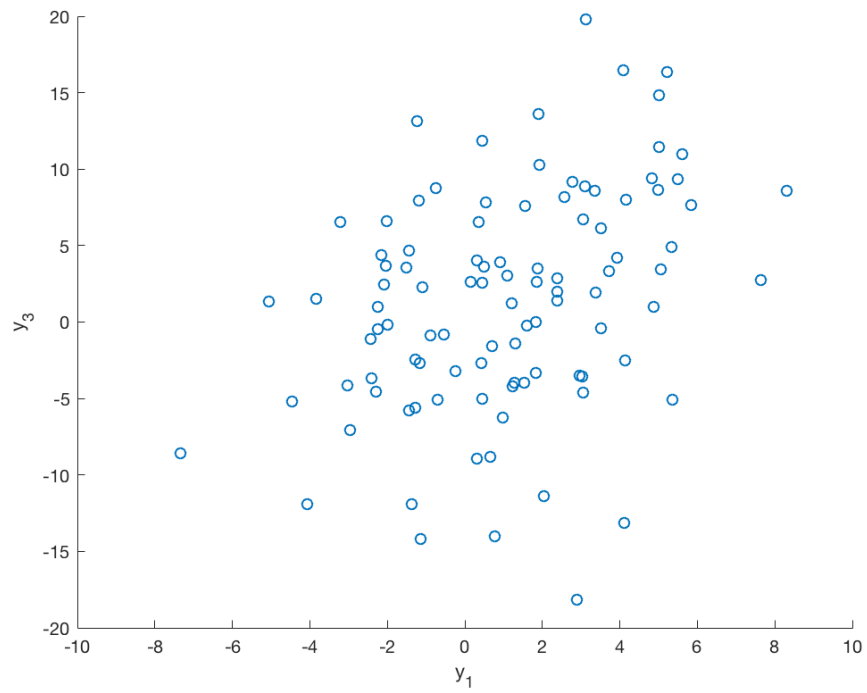
Figure 6: y_1 vs y_3
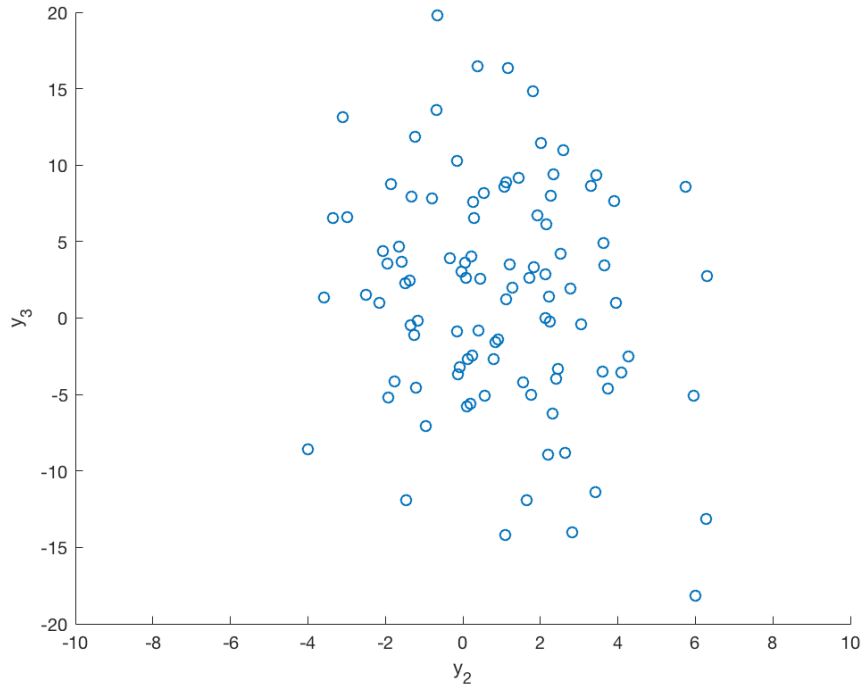
Figure 7: y_2 vs y_3

$$\text{cov}(y) = \begin{bmatrix} 8.5824 & 5.5126 & 7.8048 \\ 5.5126 & 5.1172 & -3.1177 \\ 7.8048 & -3.1177 & 53.3990 \end{bmatrix}$$

The diagonal elements of the sample covariance matrix are within 8% of their corresponding values in R. The off diagonal elements are less accurate and are within 24% of their corresponding values in R.

part (c)

Use the simulated data from part (a) to estimate the robot's 3D position via batch weighted least squares using the first 3 samples, then the first 10 samples, and finally using all 100 samples. Report the **estimation error covariance** matrix in all 3 cases (be sure to provide units). How accurate are these estimates compared to the ground truth?

Estimates for $\hat{x}$ are shown below in table .

num samples	$\hat{x}_1$ (m)	$\hat{x}_2$ (m)	$\hat{x}_3$ (m)
3	2.094	1.531	1.607
10	2.778	2.376	1.247
100	1.091	0.95	1.43

The covariances for each estimate are shown below. Units are in meters squared. The covariances suggest that the estimates increase in accuracy as more measurements are incorporated into the

estimation. The estimates of $\hat{x}$ reflect this, as they come closer to the groundtruth as the number of measurements increases.

$$P_3 = \begin{bmatrix} 2.67 & 1.72 & 2.17 \\ 1.72 & 1.67 & -1.36 \\ 2.16 & -1.35 & 16.7 \end{bmatrix} \quad P_{10} = \begin{bmatrix} 0.800 & 0.515 & 0.650 \\ 0.515 & 0.500 & -0.407 \\ 0.650 & -0.407 & 5.00 \end{bmatrix} \quad P_{100} = \begin{bmatrix} 0.080 & 0.052 & 0.065 \\ 0.051 & 0.50 & -0.041 \\ 0.065 & -0.041 & 0.500 \end{bmatrix}$$

part (d)

Using the data log of $y(k)$ vectors over $T' = 30$ time steps in the `hw6problem5data.csv` file provided on D2L, estimate the robot's position using batch weighted least squares and report the estimated error covariance matrix (be sure to provide units).

The estimated x and measurement covariance are shown below. The units of x are in meters and covariance is in meters squared.

$$\hat{x}_{LS} = \begin{bmatrix} 4.39 \\ -16.5 \\ 42.2 \end{bmatrix} \quad \begin{bmatrix} 0.267 & 0.172 & 0.217 \\ 0.172 & 0.167 & -0.136 \\ 0.217 & -0.136 & 1.67 \end{bmatrix}$$
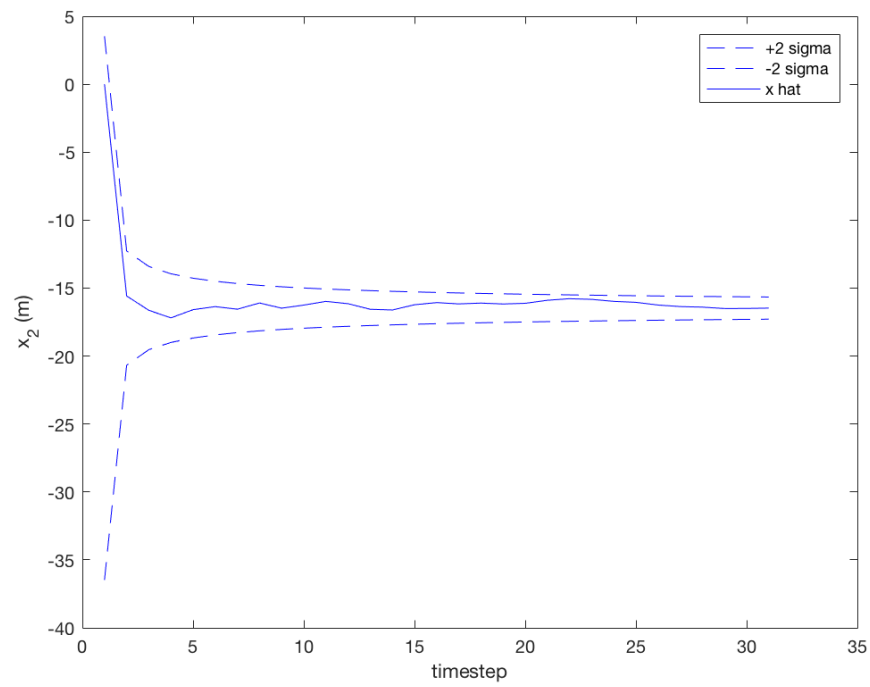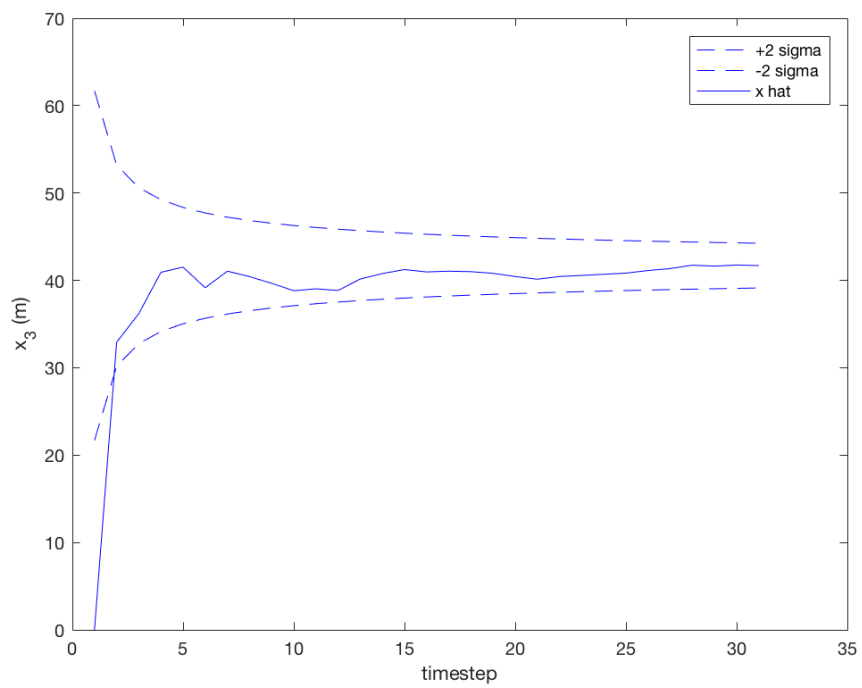
part (e)

Repeat (d) using batch unweighted least squares. Compare the final result and its covariance to the results from part (d).

The unweighted estimate is calculated as $\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} y$, except with R replaced by the identity matrix. In this case the unweighted $\hat{x}$ is actually equal to the weighted estimate. This is because the measurement covariance is constant over time. So each $y_i(k)$, the ith component of y at timestep k is weighted equally to every other y_i. The covariance of the unweighted estimate depends on the assumed variance of each measurement. If the variance is assumed to be 1 the covariance is simply $(1/n) * I$ where n is the number of measurements used.

part (f)

Repeat (d) using recursive weighted least squares and report the results by showing a separate plot for each of the estimated states (using solid lines) and their associated $\pm 2\sigma$ error bounds (using dashed lines) evolving vs time step k. In other words, plot how the state estimate and error bounds evolve as each new measurement vector $y(k)$ comes in for $k = 1$ to T' and be sure to provide units.
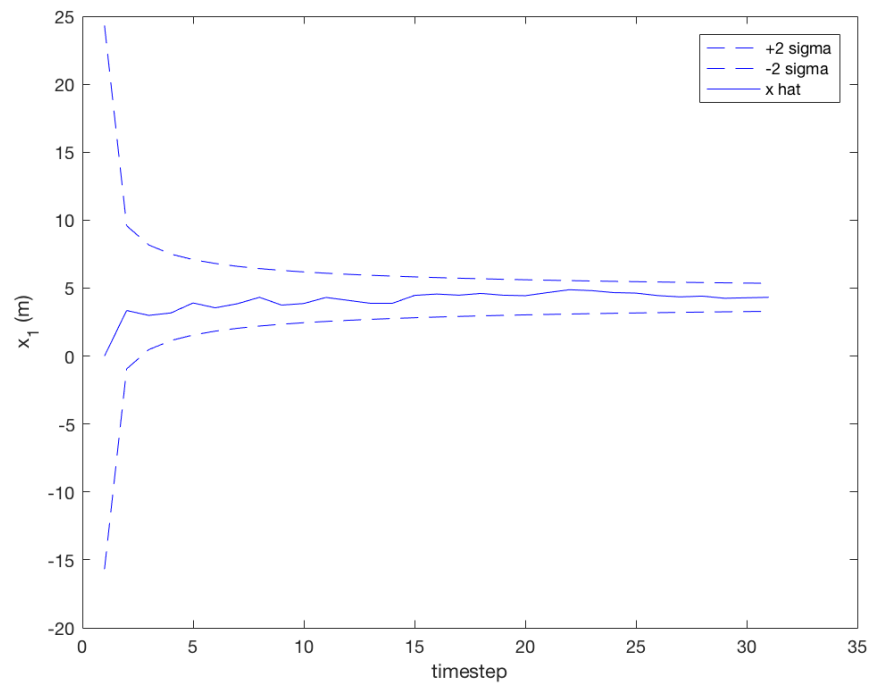
Figure 8: x_1 vs t

Figure 9: x_2 vs t

Figure 10: x_3 vs t